

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 2, 2002, 20:34:50 ; Search time 23 Seconds
(without alignments)
2814.366 Million cell updates/sec

Title: US-09-896-791B-2
Perfect score: 2010
Sequence: 1 gaattcgacagggccat.....aaaaaaaaaacatcgggccgc 1100

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09896791/runat_02122002_152659_5635/app_query.fasta_1.1287
-DB=Issued Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896791@cgn2_1_9 -runat_02122002_152659_5635 -NCPU=6 -ICPU=3
-NO_XLIFY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607.5	30.2	810	1	US-08-785-241-7
2	606.5	30.2	245	4	US-09-438-833-3
3	606.5	30.2	330	4	US-09-438-833-4
4	606.5	30.2	652	4	US-09-438-833-5
5	606.5	30.2	813	4	US-09-438-833-12
6	606.5	30.2	826	1	US-08-785-241-6
7	606.5	30.2	826	2	US-08-480-473B-2
8	606.5	30.2	826	3	US-08-915-213-2
9	606.5	30.2	826	3	US-09-148-547-2
10	606.5	30.2	826	4	US-09-235-217-2
11	606.5	30.2	826	4	US-09-380-662-23
12	606.5	30.2	826	4	US-09-438-833-1

13	606.5	30.2	826	5	PCT-US96-10251-2	Sequence 2, Appl
14	583	29.0	875	1	US-08-785-241-5	Sequence 5, Appl
15	571.5	28.4	485	4	US-09-374-454-4	Sequence 4, Appl
16	571.5	28.4	870	1	US-08-785-241-4	Sequence 4, Appl
17	571.5	28.4	870	4	US-09-374-454-6	Sequence 6, Appl
18	555.5	27.6	373	2	US-08-480-473B-3	Sequence 3, Appl
19	555.5	27.6	373	3	US-08-915-213-3	Sequence 3, Appl
20	555.5	27.6	373	4	US-09-235-217-3	Sequence 3, Appl
21	555.5	27.6	373	5	PCT-US96-10251-3	Sequence 3, Appl
22	555.5	27.6	805	2	US-08-480-473B-4	Sequence 4, Appl
23	555.5	27.6	805	3	US-08-915-213-4	Sequence 4, Appl
24	555.5	27.6	805	5	PCT-US96-10251-4	Sequence 4, Appl
25	555.5	27.6	805	5	PCT-US96-10251-4	Sequence 4, Appl
26	412	22.7	613	4	US-09-438-833-6	Sequence 6, Appl
27	412	20.5	756	4	US-09-438-833-11	Sequence 11, Appl
28	340.5	16.9	590	2	US-08-785-310A-5	Sequence 5, Appl
29	337	16.8	594	2	US-08-785-310A-6	Sequence 6, Appl
30	214	10.6	74	4	US-09-438-833-2	Sequence 2, Appl
31	203	10.1	54	2	US-08-480-473B-36	Sequence 36, Appl
32	203	10.1	54	3	US-08-915-213-36	Sequence 36, Appl
33	203	10.1	54	4	US-09-235-217-36	Sequence 36, Appl
34	199	9.9	54	4	US-09-374-454-3	Sequence 3, Appl
35	197	9.8	50	4	US-09-150-460B-17	Sequence 17, Appl
36	188	9.4	50	4	US-09-150-460B-18	Sequence 18, Appl
37	156.5	7.8	626	4	US-08-971-188-10	Sequence 10, Appl
38	156.5	7.8	789	4	US-08-971-188-9	Sequence 9, Appl
39	156.5	7.8	789	4	US-09-374-454-22	Sequence 22, Appl
40	154.5	7.7	626	4	US-09-374-454-21	Sequence 21, Appl
41	153	7.6	602	4	US-09-374-454-19	Sequence 19, Appl
42	145	7.2	808	4	US-08-971-188-12	Sequence 12, Appl
43	145	7.2	848	1	US-08-045-806-4	Sequence 4, Appl
44	145	7.2	848	1	US-08-366-051B-4	Sequence 4, Appl
45	142.5	7.1	51	4	US-09-150-460B-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-785-241-7
; Sequence 7, Application US/08785241
; Patent No. 5695963
; GENERAL INFORMATION:
; APPLICANT: Mcknight, Steven L.
; APPLICANT: Tian, Hui
; TITLE OF INVENTION: Endothelial PAS Domain Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,241
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1229
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 810 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-785-241-7

```

Alignment Scores:

Pred. No.:	9.71e-52	Length:	810
Score:	607.50	Matches:	128
Percent Similarity:	67.65%	Conservative:	33
Best Local Similarity:	53.78%	Mismatches:	49
Query Match:	30.22%	Indels:	28
Diffs:	1	Gaps:	3

US:09-896-791B-2 (1-1100) X US-08-785-241-7 (1-810)

49	QY	ACACGAGCTCGCGAAGGAGAAGTCGCGGAGCGCGCGCGAGCCGCGCAGCCAGGAG	108
		:::::	
Db	2	SerSerGluArgLysGluLysSerArgAspAlaAlaArgSerArgArgThrLysGlu	21
		:::::	
QY	109	ACGGAGTGTGTACCAAGCTGGCGGCACACTCTCCCTTTTGGCGCGCGCTCAGCGCGAC	168
		:::::	
Db	22	SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisasnValSerSerHis	41
		:::::	
QY	169	CTGGACAAGCGCTCCATCATGTGCGCCCTCAACAATCAGCTACTCTGGCATCCGCTCTGC	228
		:::::	
Db	42	LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu	61
		:::::	
QY	229	GCAGCAGGTGGAAAAGGGGAGAG-----CCACTGGAGCGCTGCTACCT	273
		:::::	
Db	62	AspAlaGlyGlyLeuAspSerGluAspGluMetLysAlaGlnMetAspCysPheTyrIle	81
		:::::	
QY	274	GAAGGGCGCTGGAGGGTTTCGTCAATGCTACTCACCGCGAGGGAGACATGGCTTACTGTC	333
		:::::	
Db	81	uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetValTyrIleSe	101
		:::::	
QY	334	GGAAATGTCAAGACACCTGGCGCTCAGTCAGTCGGACCTCTGTTCTCTCCCTCGATG	393
		:::::	
Db	101	raspAsnValasnLysTyrMetGlyLeuThrGln-----	112
		:::::	
QY	394	CATAAACCCCACTCCTGTGTACCAATTTCTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT	453
		:::::	
Db	113	-----Phe--GluLeuAlaGlyHisSerValPheasp	122
		:::::	
QY	454	TTTATCCATCCCTGTGACCAGAGGAACCTCAAGAGCGCCCTGACCCCGAGCGCGAACCTG	513
		:::::	
Db	123	PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyPro	142
		:::::	
QY	514	TCAAAGAAAGCTGGAAGCCCCAACACAGAGCGGCACCTTTTCCCTCGGAATGAAGACAG	573
		:::::	
Db	143	ValArgLysGlyLysGluLeuAsnThrGlnArgSerPhePheLeuArgMetLysCysThr	162
		:::::	
QY	574	CTCACCAGCAGAGGGCGCAGCTCAACCTCAAGCGGCGCACCTGGGAAGGTGCTGCATGCG	633
		:::::	
Db	163	LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTriplysValLeuHisCys	182
		:::::	
QY	634	TCAGGACATATGAGGGCTTACAAGCCCCCTGACACACACTCCCTCGCGGAGCCCTCGC	693
		:::::	
Db	183	ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr	200
		:::::	
QY	694	TCCGAGCGCTCCCTGCAATGCCTGGTGCCTATCTGTGAAGCCATCCCC	741
		:::::	
Db	201	LysLysProPheMetThrCysLeuValLeuIleCysGluProIlePro	216
		:::::	

RESULT 2

US-09-438-833-3

```

; Sequence 3, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: protein variant
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/

```

```

; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3
; LENGTH: 245

```

Alignment scores:

Pred. No.:	7,56e-52	Length:	245
Score:	606.50	Matches:	129
Percent Similarity:	68.0%	Conservative:	33
Best Local Similarity:	54.20%	Mismatches:	48
Query Match:	30.17%	Indels:	28
DB:	4	Gaps:	3

US-09-896-791B-2 (1-1100) x US-09-438-833-3 (1-245)

49	QY	ACACACGAGCTGCGGAAGAGAAAGTCCGGGACCGCGGCCGCGGACCGCGGACCGAGGAC	108
		:::::	
14	Db	SetSerGluArgArgLysGluArgSerArgAspAlaAlaArgSerArgSerLysGlu	33
		:::::	
109	QY	ACGGAGGTGCTGACCAGCTGGCGGCACACTCTGCCCTTTGGCGCGCGCTGCAGCGGCAC	168
		:::::	
34	Db	SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis	53
		:::::	
169	QY	CTGGACAAGCCCTCCATCATGCGCCTCACAACTACGTACTCTGGCATGCCCGCTCTGC	228
		:::::	
54	Db	LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu	73
		:::::	
229	QY	GCAGCAGGTGGAAAAAGGGGAGAG-----CCACTCGACGCGCTGCTACCT	273
		:::::	
74	Db	AspAlaGlyAspLeuAspIleGluAspAspMetLysAlaGlnMetAsnCysPheTyrIle	93
		:::::	
274	QY	GAAGGCCCTGGAGGGTTTCGTCACTACTACCGCGCGGAGGAGACATGGCTTACTGTGC	333
		:::::	
93	Db	uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe	113
		:::::	
334	QY	GGAATATGTCAGAACCACTGGCGCCTCAGTCAGTGGACCTCTGTTCCTCCTCCCTGATA	393
		:::::	
113	Db	AspAsnValAsnLysTyrMetGlyLeuThrGln-----	124
		:::::	
394	QY	CATAAACCCCACTCCTCGTACCAATTTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT	453
		:::::	
125	Db	-----Phe--GluLeuThrGlyHisSerValPheAsp	134
		:::::	
454	QY	TTTATCATCCCTGTGACCAAGAGGAACCTCAAGAGCGCCTGACCCCGGCGGACCTGC	513
		:::::	
135	Db	PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisAsnGlyLeu	154
		:::::	
514	QY	TCAAGAAGAAGCTGCAAGCCCCAACACAGAGCGCCACTTTTCCCTGCGAATGAAGACAG	573
		:::::	
155	Db	ValLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr	174
		:::::	
574	QY	CTCACGACGAGGGCGCAGCTCAACCTCAAAGCGGCCCACTGGGAAGGTGCTGCACCTGC	633
		:::::	
175	Db	LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys	194
		:::::	
634	QY	TCAGGACATATGAGGCCCTCAAGCCCCCTGCGACAGACTTCCCTCCTGCCGGGACCCCTGC	693
		:::::	
195	Db	ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr	212
		:::::	
694	QY	TCGAGCCCTCCCTGCAATGCCTGGTGCCTTATCTGTGAAGCCATCCCC	741
		:::::	
213	Db	LvsLysProMetThrCysLeuValLeuIleCysGluProIlePro	228
		:::::	

RESULT 3

US-09-438-833-4

; Sequence 4, Application US/09438833

```

; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 1-330 Of human HIF-1 alpha
US-09-438-833-4

Alignment Scores:
Pred. No.:      8,51e-52      Length:      330
Score:          606.50      Matches:    129
Percent Similarity: 68.07%      Conservative: 33
Best Local Similarity: 54.20%      Mismatches: 48
Query Match:    30.17%      Indels:     28
DB:             4          Gaps:         3

US-09-896-791B-2 (1-1100) x US-09-438-833-4 (1-330)
QY 49 ACACCGAGCTCGGAGGAGAGTCCGGGAGCGGCGGCGGCGGCGGCGGAG 108
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGGAGGTCTCTACAGCTGCGGCACACTCTGCGCTTTGCGCGCGGCTCAGCGGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 53
QY 169 CTGGCAAGGCTCCATCATGCGCTCAACAATCAGCTACCTCGCATCGCAGCGCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
QY 229 GCAGCAGGTGGAAAGGGGGAGAG-----CCACTGGAGCGCTGTACTCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGCGCTGGAGGTTTCGTCTGCTACTCACCAGGAGGAGGAGACATCGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATAATGTCAAGCAGCACCTGGGCTCAGTCAGTGGACCTCTGTCTCTCCCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAACCCCATCTCTGGTACCAATTCTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCTGTGACCAAGAGAACTTCAAGAGCGCTTGACCCCGGCGGCGGAGACCTGTC 513
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAAGCTGGAAGCCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGAGCAGG 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACCGAGAGGGCGGCGGCTCAACCTCAAGCGGCGGCGGCGGCGGAGGTGTGCTGCTGC 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATAGGGCCCTACAAGCCCCCTCGCACAGACTTCCCTCCCGGAGCCCTCCG 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCCTCCCTGCAATGCCTGTGCTTATCTGTGAAGCCATCCCC 741

```

```

Db 213 LysLysProMetThrCysLeuValLeuIleCysGluProIlePro 228
RESULT 4
US-09-438-833-5
; Sequence 5, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 1-652 Of human HIF-1 alpha
US-09-438-833-5

Alignment Scores:
Pred. No.:      1.12e-51      Length:      652
Score:          606.50      Matches:    129
Percent Similarity: 68.07%      Conservative: 33
Best Local Similarity: 54.20%      Mismatches: 48
Query Match:    30.17%      Indels:     28
DB:             4          Gaps:         3

US-09-896-791B-2 (1-1100) x US-09-438-833-5 (1-652)
QY 49 ACACCGAGCTCGGAGGAGAGTCCGGGAGCGGCGGCGGCGGCGGCGGAG 108
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGGAGGTCTCTACAGCTGCGGCACACTCTGCGCTTTGCGCGCGGCTCAGCGGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 53
QY 169 CTGGCAAGGCTCCATCATGCGCTCAACAATCAGCTACCTCGCATCGCAGCGCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
QY 229 GCAGCAGGTGGAAAGGGGGAGAG-----CCACTGGAGCGCTGTACTCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGCGCTGGAGGTTTCGTCTGCTACTCACCAGGAGGAGGAGACATCGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATAATGTCAAGCAGCACCTGGGCTCAGTCAGTGGACCTCTGTCTCTCCCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAACCCCATCTCTGGTACCAATTCTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCTGTGACCAAGAGAACTTCAAGAGCGCTTGACCCCGGCGGCGGAGACCTGTC 513
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAAGCTGGAAGCCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGAGCAGG 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACCGAGAGGGCGGCGGCTCAACCTCAAGCGGCGGCGGCGGCGGAGGTGTGCTGCTGC 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194

```

```
QY 634 TCAGGACATATGAGGGCTACAAAGCCCTGCACAGACTTCCCTGCGCGGAGCCCTCGC 693
    :|||||:
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
    :|||||:
QY 694 TCCGAGCTCCCTGCAGTCCCTGGTCTTATCTGTGAAGCCATCCCC 741
    :|||||:
Db 213 LysLysProMetThrCysLeuValLeuIleCysGluProIlePro 228
    :|||||:
RESULT 5
US-09-438-833-12
; Sequence 12, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 1-813 of human HIF-1 alpha
US-09-438-833-12

Alignment Scores:
Pred. No.: 1,22e-51 Length: 813
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 4 Gaps: 3

US-09-896-791B-2 (1-1100) x US-09-438-833-12 (1-813)
QY 49 AACACCGAGTCGGAGGAGAGTCCGGGAGCGCGCGCCGCGGCGCAGCCAGGAG 108
    :|||||:
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
    :|||||:
QY 109 ACGGAGTGTGTACACAGTCGGCGCACACTCTGCCCTTTCCGCGCGCGCTCAGCGGCAC 168
    :|||||:
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisValSerSerHis 53
    :|||||:
QY 169 CTGACAGGCTCCATCATGCGCCTCACAAATCAGCTACTGCGCATGCGCCCTCTGC 228
    :|||||:
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
    :|||||:
QY 229 GCACGAGTGGAAAAGGGGAGAG-----CCACTGGAGCCCTGTACCT 273
    :|||||:
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
    :|||||:
QY 274 GAAGGCTCGGAGGTTTCGTCTGTTACATCACCAGCGGAGGAGACATGCTTACTGTC 333
    :|||||:
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
    :|||||:
QY 334 GGAATATGTGACGAAGCACTGGCGCTCAGTCAGTGGAGACCTCTGTTCTCTCCCTCGATA 393
    :|||||:
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
    :|||||:
QY 394 CATAACCCCACTCTGTTACCATTTCTCTCTGAGCTCATTTGGACACAGTATCTTTGAT 453
    :|||||:
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
    :|||||:
QY 454 TTTATCCATCCCTGTGACCAAGAGGAACCTTCAAGACGCTGACCCCGGCGGACCTG 513
    :|||||:
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
    :|||||:
QY 514 TCAAGAAGAAGCTGGAAGGCCCCAACAGAGCGCCACTTTTCCCTGGGAATGAAGACAG 573
    :|||||:
```

```
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
    :|||||:
QY 574 CTCACGACGAGAGCGCGCAGCTCAACCTCAAGCGCGCCACCTGGAAGGTGCTGCACTGC 633
    :|||||:
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
    :|||||:
QY 634 TCAGGACATATGAGGGCTACAAAGCCCTGCACAGACTTCCCTGCGCGGAGCCCTCGC 693
    :|||||:
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
    :|||||:
QY 694 TCCGAGCTCCCTGCAGTCCCTGGTCTTATCTGTGAAGCCATCCCC 741
    :|||||:
Db 213 LysLysProMetThrCysLeuValLeuIleCysGluProIlePro 228
    :|||||:
RESULT 6
US-08-785-241-6
; Sequence 6, Application US/08785241
; Patent No. 5695963
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; APPLICANT: Tian, Hui
; TITLE OF INVENTION: Endothelial PAS Domain Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,241
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1229
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-785-241-6

Alignment Scores:
Pred. No.: 1,23e-51 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 1 Gaps: 3

US-09-896-791B-2 (1-1100) x US-08-785-241-6 (1-826)
QY 49 AACACCGAGTCGGAGGAGAGTCCGGGAGCGCGCGCCGCGGCGCAGCCAGGAG 108
    :|||||:
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
    :|||||:
QY 109 ACGGAGTGTGTACACAGTCGGCGCACACTCTGCCCTTTCCGCGCGCGCTCAGCGGCAC 168
    :|||||:
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisValSerSerHis 53
    :|||||:
QY 169 CTGACAGGCTCCATCATGCGCCTCACAAATCAGCTACTGCGCATGCGCCCTCTGC 228
    :|||||:
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
    :|||||:
QY 229 GCACGAGTGGAAAAGGGGAGAG-----CCACTGGAGCCCTGTACCT 273
    :|||||:
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
    :|||||:
QY 274 GAAGGCTCGGAGGTTTCGTCTGTTACATCACCAGCGGAGGAGACATGCTTACTGTC 333
    :|||||:
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
    :|||||:
QY 334 GGAATATGTGACGAAGCACTGGCGCTCAGTCAGTGGAGACCTCTGTTCTCTCCCTCGATA 393
    :|||||:
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
    :|||||:
QY 394 CATAACCCCACTCTGTTACCATTTCTCTCTGAGCTCATTTGGACACAGTATCTTTGAT 453
    :|||||:
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
    :|||||:
QY 454 TTTATCCATCCCTGTGACCAAGAGGAACCTTCAAGACGCTGACCCCGGCGGACCTG 513
    :|||||:
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
    :|||||:
QY 514 TCAAGAAGAAGCTGGAAGGCCCCAACAGAGCGCCACTTTTCCCTGGGAATGAAGACAG 573
    :|||||:
```



```

Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 53
QY 169 CTGGACAGGCTCCATCATCGCCCTCAACAATCAGTACTCGGATGCACCGCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgLysLeuLeu 73
QY 229 GCAGCAGGTGGAAAGGGGAGAG-----CCACTGGAGCGCTCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCGCTGGAGGTTTCGTATGTTACTACCGCGGAGGAGACATGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIle 113
QY 334 GGAATATGTCAGCAAGCAGCTGGGCTCAGTCAGTGGACCTCTGTCTCCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAACCCCTCTCTGGTACCAATTTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCCTGTGACCAAGAGGAACTTCAAGAGCGCCCTGACCCCGGAGGAGACCTG 513
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGCAGTGGAGCGCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGACGAG 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACGAGCAGGCGCGCACCGCTCAACCTCAAGCGCCGAGGAGGAGGAGGAGGAGGAGG 633
Db 175 LeuThrSerArgLysGlyThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGCGCTTACAGCGCCCTGACAGACTTCCCTCGCGGAGCGCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCTCCCTCGCAATGCTGTGCTTATCTGTGAAGCCATCCCC 741
Db 213 LysLysProMetThrCysLeuValLeuIleCysGluProIlePro 228

```

```

RESULT 7
US-08-480-473B-2
; Sequence 2, Application US/08480473B
; Patent No. 5882914
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,473B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-473B-2
Alignment Scores:
Pred. No.: 1,23e-51 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 2 Gaps: 3
US-09-896-791B-2 (1-1100) x US-08-480-473B-2 (1-826)
QY 49 AACACCGAGCTGCGGAAGGAGAGTCCGCGGACGCGCGCGCGCGCGCGCGCGAG 108
Db 14 SerSerGluArgGlyLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGGAGGTGCTGTACAGCTGGCGGCACACTCTGCCCTTTGCGCGCGCGCTCAGCGCGAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 53
QY 169 CTGGACAGGCTCCATCATCGCCCTCAACAATCAGTACTCGGATGCACCGCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgLysLeuLeu 73
QY 229 GCAGCAGGTGGAAAGGGGAGAG-----CCACTGGAGCGCTCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCGCTGGAGGTTTCGTATGTTACTACCGCGGAGGAGACATGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIle 113
QY 334 GGAATATGTCAGCAAGCAGCTGGGCTCAGTCAGTGGACCTCTGTCTCCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAACCCCTCTCTGGTACCAATTTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCCTGTGACCAAGAGGAACTTCAAGAGCGCCCTGACCCCGGAGGAGACCTG 513
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGCAGTGGAGCGCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGACGAG 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACGAGCAGGCGCGCACCGCTCAACCTCAAGCGCCGAGGAGGAGGAGGAGGAGGAGG 633
Db 175 LeuThrSerArgLysGlyThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGCGCTTACAGCGCCCTGACAGACTTCCCTCGCGGAGCGCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCTCCCTCGCAATGCTGTGCTTATCTGTGAAGCCATCCCC 741
Db 213 LysLysProMetThrCysLeuValLeuIleCysGluProIlePro 228

```

```

RESULT 8
US-08-915-213-2
; Sequence 2, Application US/08915213
; Patent No. 6020462
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE

```

NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,213
FILING DATE: 20-AUG-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,473
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/053001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 826 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-915-213-2

Alignment Scores:
Pred. No.: 1,23e-51 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 3 Gaps: 3

US-09-896-791b-2 (1-1100) x US-08-915-213-2 (1-826)

```
QY 49 AACACCGAGCTGCGGAGGAGAGTCCGCGGACGCGCGCGCGCGCGCGCGCGCGCGAG 108
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGGAGGTGCTGACCGAGTGGCGCACACTCTGCCCTTTTCGCGCGCGCGCGCGCGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerHis 53
QY 169 CTGACGAAGCGCTCCATCATCGCGCTCACAAATCAGCTACCTGCGCATGCGCGCGCTGCG 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeu 73
QY 229 GCAGCAGTGGAAAAAGGGGAGAG-----CCACTGGACCGCTGCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCCCTGGAGGGTTTCGTCATGTGTACTACCGCGGAGGAGACATGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTACGAAGCAACCTGGCGCTCAGTCAGTGAGCCTCTGTCCTCCCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAAACCCCACTCCTGGTACCAATTTCTCTCTGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
```

```
QY 454 TTTATCATCCCTGTGTGACCAGAGGAACCTTCAAGACGCCCTGTAGCCCGGCGGCGGAGCTG 513
Db 135 PheThrHisProCysAspHisGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAGAAGCTGGAGGCCCAACAGAGCGGCACCTTTTCCCTGCGGAATGAAGAGCAGC 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPheLeuArgMetLysCysThr 174
QY 574 CTCACGACGAGGCGCGCACGCTCAACCTCAAAGCGGCCACCTGGAAGGTGCTGCTGCTG 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATAGGCGCTACAAAGCCCGCTGCACAGACTTCCCTGCGCGGAGCCCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCTCCCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
Db 213 LysLysProMetThrCysLeuValLeuLeuValLeuValLeuValLeuValLeuValLeu 228
RESULT 9
US-09-148-547-2
; Sequence 2, Application US/09148547
; Patent No. 6124131
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: Hypoxia Inducible Factor-1 and Methods of Use
; FILE REFERENCE: 07265/151001
; CURRENT APPLICATION NUMBER: US/09/148,547
; CURRENT FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-148-547-2
Alignment Scores:
Pred. No.: 1,23e-51 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 3 Gaps: 3
US-09-896-791b-2 (1-1100) x US-09-148-547-2 (1-826)
QY 49 AACACCGAGCTGCGGAGGAGAGTCCGCGGACGCGCGCGCGCGCGCGCGCGCGAG 108
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGGAGGTGCTGACCGAGTGGCGCACACTCTGCCCTTTTCGCGCGCGCGCGCGCGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerHis 53
QY 169 CTGACGAAGCGCTCCATCATCGCGCTCACAAATCAGCTACCTGCGCATGCGCGCGCTGCG 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeu 73
QY 229 GCAGCAGTGGAAAAAGGGGAGAG-----CCACTGGACCGCTGCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCCCTGGAGGGTTTCGTCATGTGTACTACCGCGGAGGAGACATGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTACGAAGCAACCTGGCGCTCAGTCAGTGAGCCTCTGTCCTCCCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAAACCCCACTCCTGGTACCAATTTCTCTCTGAGCTCATTTGGACACAGTATCTTTGAT 453
```

[illegible][illegible]

```

Qy 594 TCCGAGCCTCCCTCGCAATGCCCTGGGTGCTTATCTGTGAGACCATCCCC 741
      :|||||:  |||||:|||||:|||||:|||||:|||||:
Db 213 LysLysProMetThrCysLeuValLeuileCysgluProilePro 228

RESULT 11
US-09-380-662-23
; Sequence 23, Application US/09380662
; Patent No. 6376199
; GENERAL INFORMATION:
; APPLICANT: Caniggia, Isabella
; APPLICANT: Post, Martin
; APPLICANT: Lye, Stephen
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF TROPHO
; FILE REFERENCE: 11757.38USWO
; CURRENT APPLICATION NUMBER: US/09/380,662
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; PRIOR FILING DATE: 1998-03-05

```

```

; PRIOR APPLICATION NUMBER: US 60/039,919
;
; PRIOR FILING DATE: 1997-03-07
;
; NUMBER OF SEQ ID NOS: 24
;
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 23
;
; LENGTH: 826
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
US-09-380-662-23

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 2, 2002, 20:39:21 : Search time 40 Seconds
(without alignments)
875.831 Million cell updates/sec

Title: US-09-896-791B-2
Perfect score: 2010
Sequence: 1 gaattcgacagggccat.....aaaaaaaaaacatcgggcgc 1100

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 204634

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09896791/runat_02122002_152700_5721/app_query.fasta_1.1287
-DB=PublishedApplications_AA -QPMF=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09896791@cgn_1.1_5@runat_02122002_152700_5721
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1636	81.4	307	9	US-09-896-791B-3
2	606.5	30.2	826	9	US-09-736-457-330
3	606.5	30.2	826	9	US-09-902-941-330
4	606.5	30.2	826	10	US-09-922-958-4

5	606.5	30.2	826	10	US-09-833-790-235
6	606.5	30.2	826	12	US-10-028-158-23
7	311	15.5	570	10	US-09-923-684-4
8	311	15.5	667	10	US-09-923-684-3
9	222.5	11.1	99	10	US-09-925-297-863
10	192	9.6	333	10	US-09-925-302-659
11	162	8.1	591	10	US-09-764-864-1139
12	153	7.6	53	10	US-09-864-761-41589
13	144.5	7.2	530	10	US-09-800-729-112
14	134	6.7	399	9	US-09-764-868-850
15	132.5	6.7	200	10	US-09-764-864-953
16	128.5	6.4	1464	10	US-09-842-256-2
17	127.5	6.3	223	10	US-09-925-301-1280
18	126	6.3	181	10	US-09-764-846-264
19	123	6.3	178	9	US-09-764-868-650
20	120.5	6.0	292	10	US-09-764-853-604
21	119.5	5.9	315	10	US-09-764-853-793
22	119	5.9	219	12	US-10-001-843-143
23	116.5	5.8	640	12	US-10-028-780-2
24	115	5.9	60	10	US-09-764-846-178
25	115	5.7	359	10	US-09-864-761-36539
26	115	5.7	846	10	US-09-858-684A-3
27	114.5	5.7	133	10	US-09-925-301-1086
28	113	5.6	201	10	US-09-764-864-1561
29	111.5	5.5	881	10	US-09-816-860A-2
30	111	5.7	51	10	US-09-764-846-151
31	111	5.5	154	10	US-09-825-301-27
32	111	5.5	154	10	US-09-604-287A-471
33	111	5.5	154	12	US-10-007-805-471
34	110.5	5.5	677	10	US-09-925-300-1626
35	110.5	5.5	1350	10	US-09-893-238-17
36	109	5.4	196	10	US-09-764-846-251
37	108.5	5.4	645	10	US-09-925-300-1453
38	107.5	5.3	150	10	US-09-764-846-225
39	107.5	5.3	388	10	US-09-925-300-1618
40	107.5	5.3	1317	10	US-09-963-896-7
41	107	5.4	80	10	US-09-764-846-158
42	106.5	5.3	1429	10	US-09-996-617-2
43	106.5	5.3	1429	10	US-09-931-071-2
44	106	5.4	50	10	US-09-764-846-186
45	106	5.4	60	10	US-09-764-846-187

ALIGNMENTS

RESULT 1
US-09-896-791B-3
; Sequence 3, Application US/09896791B
; Patent No. US20020165140A1
; GENERAL INFORMATION:
; APPLICANT: Berkenstam, Anders
; APPLICANT: Bertilsson, Gran
; APPLICANT: Poellinger, Lorenz
; TITLE OF INVENTION: SCREENING METHODS
; FILE REFERENCE: 13425-040001
; CURRENT APPLICATION NUMBER: US/09/896,791B
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/217,570
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: SE 0002551-0
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 3
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-896-791B-3

Alignment Scores: 3.93e-120 Length: 307
Pred. No.: 1636.00 Matches: 307
Score: 100.00%
Percent Similarity: 100.00%
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 81.39% Indels: 0
DB: 9 Gaps: 0

US-09-896-791B-2 (1-1100) x US-09-896-791B-3 (1-307)

```
QY 19 ATGCGTTGGGCTGCGCGCGTGAAGTGCAGACCGAGCTGCGGAAGAGAGTCCGG 78
Db 1 MetAlaLeuGlyLeuGlnArgValArgSerAsnThrGluLeuArgLysGluLysSerArg 20
QY 79 GACCGCGCGCGCAGCGCGCGCAGCGAGGAGGAGGTGCTGACAGCTGGCGCACACT 138
Db 21 AspAlaalaArgSerArgSerGlnGlnThrGluValLeuLeuGlnLeuAlaHisThr 40
QY 139 CTGCGCTTTGCGCGCGCGTGCAGCGCGCAGTGCAGACAGGCTCCATCATCGCGCTCACA 198
Db 41 LeuProPheAlaArgGlyValSerAlaHisLeuAspLysAlaSerIleMetArgLeuThr 60
QY 199 ATCAGCTACCTGCGCATGCACCGCTCTGCGCAGCAGGTGGAAGGAGGAGAGCCACT 258
Db 61 IleSerTyrLeuArgMetHisArgLeuCysAlaAlaGlyGlyLysArgGlyArgAlaThr 80
QY 259 GGACCGCTGTACTGAAGCGCTGAGGCTTCGTATGCTATGCTACCGCCCGAGGAGA 318
Db 81 GlyArgLeuLeuProGluGlyProGlyGlyPheArgHisGlyThrHisArgGlyArg 100
QY 319 CATGGCTTACCTGTCGGAATGTCAGCAGCAGCTGGCTCAGTCAGTGGACCTCTGT 378
Db 101 HisGlyLeuProValGlyLysCysGlnGlnAlaProGlyProGlnSerValAspLeuCys 120
QY 379 TCCTCTCCCTGTATACATAACCCACCTCCCTGGTACCAATTTCTCTCTGGAGCTCATTTGA 438
Db 121 SerSerSerLeuIleHisAsnProThrProGlyThrAsnPheSerLeuGluLeuIleGly 140
QY 439 CACAGTATCTTTGATTTTATCCATCCCTGTGACCAAGAGGAATTCGAAGCGCCCTGACC 498
Db 141 HisSerIlePheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThr 160
QY 499 CCAGGCGCGAACCTGTCAAAGAAGACTGGAAGCCCAACAGAGCGCCACTTTCCCTG 558
Db 161 ProArgProAsnLeuSerLysLysLysLeuGluAlaProThrGluArgHisPheSerLeu 180
QY 559 CGAATGAGAGCAGCTCACCAGCAGAGCGCGCACGCTCAACCTCAAAGCGGCCACCTGG 618
Db 181 ArgMetLysSerThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrp 200
QY 619 AAGTGCTGCTGCTCAGACATATAGAGGCGCTACAGCCCGCTGCACAGACTTCCCGCT 678
Db 201 LysValLeuHisCysSerGlyHisMetArgAlaTyrLysProAlaGlnThrSerPro 220
QY 679 GCCGGAGCCCTCGCTCCGAGCCTCCCTGCAATGCTGCTGTATCTCTGAGCCATC 738
Db 221 AlaGlySerProArgSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIle 240
QY 739 CCCAGCTCCCGCTTCCACGATGCTGCTACTCTGGGTCTTCCACAGGAGAAGACTCCCATC 798
Db 241 ProGlnLeuProPheHisAspGlyAlaThrLeuGlyLeuProGlnGluLysThrProIle 260
QY 799 TCTACCTTATACCCCTCTTTGGAAGGCACTACTTTGCTTGTTCAGAGAGGTGGCTGTT 858
Db 261 SerThrLeuPheThrProLeuTrpLysAlaLeuLeuCysLeuValLysArgTrpProVal 280
QY 859 CAGTGCTACAGGGGAAGGACTGAATCTCTCTCCCGCTCATGGGTGTTGTGGCGCTT 918
Db 281 GlnValLeuGlnGlyLysGlyThrGluSerSerLeuProSerTrpValLeuTrpAlaLeu 300
QY 919 AACCGGAAAAATGTCCTGGC 939
Db 301 AsnArgLysAsnCysProGly 307
```

RESULT 2

US-09-736-457-330

; Sequence 330, Application US/09736457

; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Hannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-736-457-330

Alignment Scores:

Pred. No.: 7,65e-40 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 9 Gaps: 3

US-09-896-791B-2 (1-1100) x US-09-736-457-330 (1-826)

```
QY 49 AACACCGAGCTGCGGAAGAGAAAGTGCAGGAGCGCGCGCGCGCGCGAGGAG 108
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGAGGTGCTGACACAGCTGGCGCACACTCTGCCCTTTCGCGCGCGCTCAGCGCGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerHis 53
QY 169 CTGGAAGCGCTTCATCATGCGCTCACAAATCAGCTACCTGCGCATGCGCGCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
QY 229 GCACGAGGTGGAAAAAGGGGAGAG-----CCACTGGAGCGCTGCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCGCTGGAGGGTTTCGTATGCTGCTACCGCGCGGAGGAGACATGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTCAGCAAGCACCTGGCGCTCAGTCAAGTGGACCTCTGCTCCTCCTCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAACCCCACTCCTGCTACCAATTTCTCTCTGGAGCTCAATTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCCTGTGACCAAGAGGAACCTTCAAGACGCCCTGACCCCGCGCGCAACCTG 513
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAAGAAGAAGCTGGAAGCCCAACAGAGCGCACCTTTTCCTCGCGAATGAAGAGACAG 573
Db 155 valLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACGAGCAGAGGCGCGCAGCTCAACCTCAAAGCGGCCACCTGGAAGGTGCTGCACTGC 633
|||||
```


Db 175 LeuThrSerArglyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGCGCTACAAAGCCCTGCACAGACTTCCCTCGCGGAGCCCTCGC 693
Db 195 ThrGlyHisIleHisValTrpAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCCTCCCTGCAATGCCGTGGTCTATCTGTGAAGCCATCCCC 741
Db 213 LysLysProPromethrCysLeuValLeuIleCysGluProIlePro 228
RESULT 3
US-09-902-941-330
; Sequence 330, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongcong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-330
Alignment Scores:
Pred. No.: 7.65e-40 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 9 Gaps: 3
US-09-896-791B-2 (1-1100) x US-09-902-941-330 (1-826)
QY 49 AACACCGAGCTGCGGAAGAGAGTCCGCGGACGCGCGCGCGCGCGCGCGCAGGAG 108
Db 14 SerSerGluArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGAGGTGCTGTACAGCTGGCGGCACACTCTGCCCTTTTCCGCGCGCGCGTACGCGGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 53
QY 169 CTGGACAAGCCCTCCATCATGCGCTCACAATCAGTACTGCGCATGACCGCCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
QY 229 GCAGCAGGTGGAAAAAGGGGAGAG-----CCACTGAGCGCTGTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetCAsnCysPheTyrLe 93
QY 274 GAAGGCCCTGGAGGTTTGGTTCATGTTACTACCGCGGAGGAGACATGGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTCACGACACCTGGGCGCTCAGTCAGTGACCTCTGTTCCTCCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124

QY 394 CATATACCCCACTCCTGGTACCAATTTCTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCATCCCTGTGACCAAGAGAACTTCAAGAGCCCTGACCCCGCAGGCGCGAACCTG 513
Db 135 PheThrHisProcysAspHisGluGluMetArgLysLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAAGCTGGAAGCCCAACAGAGCGCCACTTTTCCCTCGCAATGAAGAGCACG 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACGAGGAGGCGCGACGCTCAACCTCAAAAGCGCCACCTGGAGGTGCTGTCACCTGC 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGCGCTACAAAGCCCTGTCACAGACTTCCCTCGCGGAGCCCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCCTCCCTGCAATGCCGTGGTCTATCTGTGAAGCCATCCCC 741
Db 213 LysLysProPromethrCysLeuValLeuIleCysGluProIlePro 228
RESULT 4
US-09-922-958-4
; Sequence 4, Application US/09922958
; Patent No. US20020048794A1
; GENERAL INFORMATION:
; APPLICANT: POELLINGER, Lorenz
; APPLICANT: PEREIRA, Teresa
; APPLICANT: RUAS, Jorge
; TITLE OF INVENTION: MECHANISM OF CONDITIONAL REGULATION OF THE HYPOXIA-INDUCIBLE F
; TITLE OF INVENTION: THE VON HIPPEL-LINDAU TUMOR SUPPRESSOR PROTEIN
; FILE REFERENCE: 3743/49008
; CURRENT APPLICATION NUMBER: US/09/922,958
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/223,480
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-958-4
Alignment Scores:
Pred. No.: 7.65e-40 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 10 Gaps: 3
US-09-896-791B-2 (1-1100) x US-09-922-958-4 (1-826)
QY 49 AACACCGAGCTGCGGAAGAGAGTCCGCGGACGCGCGCGCGCGCGCGCGCAGGAG 108
Db 14 SerSerGluArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGAGGTGCTGTACAGCTGGCGGCACACTCTGCCCTTTTCCGCGCGCGCGTACGCGGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 53
QY 169 CTGGACAAGCCCTCCATCATGCGCTCACAATCAGTACTGCGCATGACCGCCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
QY 229 GCAGCAGGTGGAAAAAGGGGAGAG-----CCACTGAGCGCTGTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetCAsnCysPheTyrLe 93

```
QY 274 GAAGGCCCTCGAGGGTTTCCTCATGGTACATCGCGGAGGAGACATGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTCAGCAAGCACCTGGGCCTCAGTCAGTGCACCTCTGTTCTCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAAACCCTCTCTGGTACCAATTTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCCTGTGACCAAGAGAACTTCAAGACGCCCTGACCCCGAGCGCAACCTG 513
Db 135 PheThrHisProCysAspHisGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAGCTGGAAGCCCAACAGAGCGGCACCTTTTCCCTGCGAATGAAGAGCAGC 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTACCAAGAGCGGCACGCTCAAAAGCGCCACCTGGAAGGTGCTGCACTGC 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGCGCTACAAGCCCGCTGCACAGACTTCCCTGCGCGGAGCCCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCCTCCCTGCAATCGCTGGTGTATCTGTGAAGCCATCCCC 741
Db 213 LysLysProMetThrCysLeuValIleCysGluProIlePro 228
```

RESULT 5

```
US-09-833-790-235
; Sequence 235, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indirias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121-512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 235
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-235
```

```
Alignment Scores:
Pred. No.: 7 65e-40 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservativity: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 10 Gaps: 3
```

US-09-896-791B-2 (1-1100) x US-09-833-790-235 (1-826)

```
QY 49 AACACGAGTGGGAGGAGAGTCCGCGGACGCGCGCCGCGCAGCGCGCAGCAGGAG 108
Db 14 SerSerGluArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGAGGTGCTACCAAGTGGCGCACACTCTGCCCTTTGCGGCGCGCTCAGCGCGCAC 168
Db 34 SerGluValPheTyrGlyLeuAlaHisGlnLeuProLeuProHisAsnValSerHis 53
```

7-6

```
QY 169 CTGCACAAGCCCTCCATCATCGCCTCACATCAGCTACCTGGCATCACCAGCCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLe 73
QY 229 GCAGCAGTGGAAAAGGGGAGAG-----CCACTGGACGCGCTGCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCCCTCGAGGGTTTCCTCATGGTACATCGCGGAGGAGACATGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTCAGCAAGCACCTGGGCCTCAGTCAGTGCACCTCTGTTCTCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAAACCCTCTCTGGTACCAATTTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCCTGTGACCAAGAGAACTTCAAGACGCCCTGACCCCGAGCGCAACCTG 513
Db 135 PheThrHisProCysAspHisGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAGCTGGAAGCCCAACAGAGCGGCACCTTTTCCCTGCGAATGAAGAGCAGC 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTACCAAGAGCGGCACGCTCAAAAGCGCCACCTGGAAGGTGCTGCACTGC 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGCGCTACAAGCCCGCTGCACAGACTTCCCTGCGCGGAGCCCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCCTCCCTGCAATCGCTGGTGTATCTGTGAAGCCATCCCC 741
Db 213 LysLysProMetThrCysLeuValIleCysGluProIlePro 228
```

RESULT 6

```
US-10-028-158-23
; Sequence 23, Application US/10028158
; Patent No. US20020110833A1
; GENERAL INFORMATION:
; APPLICANT: Caniggia, Isabella
; APPLICANT: Post, Martin
; APPLICANT: Lye, Stephen
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF
; TITLE OF INVENTION: TROPHOBLAST
; FILE REFERENCE: 11757-38USWO
; CURRENT APPLICATION NUMBER: US/10/028,158
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US/09/380,662
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-158-23
```

```
Alignment Scores:
Pred. No.: 7 65e-40 Length: 826
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservativity: 33
Best Local Similarity: 54.20% Mismatches: 48
```

Query Match:	30.17%	Indels:	28
DB:	12	Gaps:	3
 US-09-896-791B-2 (1-1100) x US-10-028-158-23 (1-826)			
QY	49	AACACCAGCTGCGGAGGAGAAGTCGGCGGACGCCGCCGACCGCCGACGCAGCAGGAG	108
Db	14	SerSerGluArgGlySGLysSerArgAspAlaAArgSerArgArgSerLysGlu	33
QY	109	ACGGAGTGCTGTACACAGCTGGCGCACACTCTCCCTTTTGGCGGCGCTCAGCGGCAC	168
Db	34	SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisasnValSerSerHis	53
QY	169	CTGGACAAGCCCTCCATCATCTGCCTCCACAAATCAGCTACTCTGGCATCGACCCCTCTGC	228
Db	54	LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu	73
QY	229	GCAGCAGGTGGAAAAGGGGGAG-----CCACTGGACGCGTGTCACT	273
Db	74	AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCynPheTyrLe	93
QY	274	GAAGGCCCTGGAGGTTTCGTGATGGTACTACCCGCCGAGGGAGACATAGCTTACCTGTC	333
Db	93	uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe	113
QY	334	GGAATAATGTCAAGAAGCACCTGGGCGCTCAGCTCAGTGGACCTCTGTTCCCTCCCTGATA	393
Db	113	rAspAsnValAsnLysTyrMetGlyLeuThrGln-----	124
QY	394	CATAACCCCCTCCTGGTACCAATTCTCTCTGGAGCTATTGGACACAGATCTTTGAT	453
Db	125	-----Phe--GluLeuThrGlyHisSerValPheAsp	134
QY	454	TTTATCATCCTGTGACCAAGAGGAACTTCAAAGCGCCTGACCCCCAGCGCGAACCTG	513
Db	135	PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgasnGlyLeu	154
QY	514	TCAAAGAAGAAGCTGGGAAGCCCCAACAGACGCGCCTTTTCCCTGCGAATGAAGACGACG	573
Db	155	ValllysGlyLysGlyGluGlnAsnThrGlnArgSerPheLeuLeuargMetLysCysThr	174
QY	574	CTCACACAGAGGGCGCAGCTCAACCTCAAAGCGGCCACCTGGGAAGGTGTGTCACCTGC	633
Db	175	LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys	194
QY	634	TCAGGACATATGAGGGCCTACAAAGCCCCCTGACAGACTTCCCTCCCGGGAGCCCTCGC	693
Db	195	ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr	212
QY	694	TCCGAGCCTCCCTGCNAATCCCTGGTGTCTATCTGTGAAGCCATCCCC	741
Db	213	LysLysProMetThrCysLeuValLeuIleCysGluProIlePro	228

```

RESULT 7
US-09-923-684-4
; Sequence 4, Application US/09923684
; Patent No. US20020081613A1
; GENERAL INFORMATION:
; APPLICANT: Narayanan, Ramaswamy
; TITLE OF INVENTION: ASSOCIATION OF SIM2 WITH CANCER
; FILE REFERENCE: 6818-24
; CURRENT APPLICATION NUMBER: US/09/923,684
; CURRENT FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-684-4

Alignment Scores:
Pred. No.: 7,336-17 Length: 570

```

Score:	311.00	Matches:	124
Percent Similarity:	42.38%	Conservative:	54
Best Local Similarity:	29.52%	Mismatches:	123
Query Match:	15.47%	Indels:	119
DB:	10	Gaps:	15
US-09-896-791B-2 (1-1100) x US-09-923-684-4 (1-570)			
Qy	64	AAGGAGAAGTCGGGGACGGCCCGCAGCGCGCGCAGCCAGCGAGCGAGGTCGCTGAC	123
Db	2	LysGluLysSerLysAsnAlaLysThrArgArgGluLysGluAsnGlyGluPheTyr	21
Qy	124	CAGCTGGCGCACACTCGCCCTTTGCCGGCGGCTCAGCGGCACCTGGACAGGCCTCC	183
Db	22	GluLeuAlaLysLeuLeuProSerAlaIleThrSerGlnLeuAspLysAlaSer	41
Qy	184	ATCATGCCCTCAACATCAGCTACCTCGCATGCCGCTCTCCGCAGCAGCT	237
Db	42	IleIleArgLeuThrThrSerTyrLeuLysMetArgAlaValPheProGluGlyLeuGly	63
Qy	238	-----GGAAGGAGGGGAGAGCC-----	261
Db	62	AspAlaTrpGlyGlnProSerArgAlaGlyProLeuAspGlyValAlaLysGluLeuGly	81
Qy	262	CGCCTGCTACCTGAAGCCCTGGAGGTTTCGTATGTTACTCACCGCGGAGGAGACAT	321
Db	82	SerHisLeu-LeuGlnThrLeuAspGlyPheValPheValValAlaLaserAspGlyLysI	101
Qy	322	GGCTTACCTGCGAAATGTCACAGCAGCCTGGGCTCAGTCAGTGGACCTCTGTTC	381
Db	101	eMetTyrIleSerGluThrAlaSerValHisLeuGlyLeuSerGln-----	116
Qy	382	TCCTCCCTGATACATACCCCACTCCTGGTACCAATTTCTCTGGAGCTCATTTGGACAC	441
Db	117	-----Val--GluLeuThrGlyAsn-----	122
Qy	442	AGTATCTTTGATTATTCATCCCTGTGCACCAAGAGAACTTCAAGACGCCCTGACCC	501
Db	123	SerIleTyrGluTyrIleHisProSerAspHisAspGluMetThrAlaValLeuThrAla	142
Qy	502	AGCGCGAACCTGTCAAAGAGAAGCTG---GAAGCCCAACAGAGCCCACTTTCCCTG	558
Db	143	HisGlnProLeuHisHisLeuLeuGlnGluTyrGluIleGluArgSerPhePheLeu	162
Qy	559	CGAATGAAGAGCAGCTCACCAGCAGAGGGCGACGCTCAACCTCAAAGCGCCACCTGG	618
Db	163	ArgMetLysCysValLeuAlaLys-----ArgAsnAlaGlyLeuThrCysSerGlyTyr	180
Qy	619	AAGTGTGTGCTGCTCAGCAGCATATGAGGCGCTTACAGGCCCTGCACAGACTCCCT	678
Db	181	LysValIleHisCysSerGlyTyrLeuLysIleArgGlnTyrMetLeuAspMetSerLeu	200
Qy	679	GCCGGGAGC-----CCTCCCTCCGAGCCCTCCCTG	708
Db	201	TyrAspSerCysTyrGlnIleValGlyLeuValAlaValGlyGlnSerLeuPro---	219
Qy	709	CAATGCCTGGTGCTATCTGTGAAGCATCCCTCCAGCTCCCTTCCACAGTGGT---	762
Db	220	-----SerAlaIleThrGluIleLysLeuTyrHisHisValHisGlyCys	232
Qy	763	-----GCTACTCTGGTCTT-----	777
Db	233	MetPheArgAlaSerLeuAspLeuLysLeuIlePheLeuAspSerArgValThrGluVal	252
Qy	778	-----CCACAG-----GAGNAGACTCCCATC-----TCT	801
Db	253	ThrGlyTyrGluProGlnAspLeuIleGluLysThrLeuTyrHisHisValHisGlyCys	272
Qy	802	ACCTATTATCCCTCTTTGGAGGCACACTACTTTGTTCTGTCTCAAGAGGTGGCTGT---	858
Db	273	AspValPheHisLeuArgTyrAlaHisHisLeuLeuValLysGlyGlnValThrThr	292
Qy	859	-----CAGGTGCTACAGGGGAAGGG-----	879

alignment scores:	
pred. no.:	7.33e-17
length:	570

```
Db 293 LysTyrTyrArgLeuLeuSerLysArgGlyGlyTrpValTrpValGlnSerTyrAlaThr 312
QY 880 -----ACTGAATCTCTCTCTCC-----TCATGGGTCTTG 909
Db 313 ValValHisAsnSerArgSerSerArgProHisCysIleValSerValAsnTyrValLeu 332
QY 910 TGGGCCCTTAACCGGAAATAATTCCTGGCTAGGAGGAGTGAAGGACATGGCCACGTA 969
Db 333 ThrGluIleGluTyrLysGlu-----LeuGlnLeu 342
QY 970 TCCTTACCCAGAAACCCCAAAATGCTCCAAAACACCAATAAAGACCTCTCTTGTGTAG 1029
Db 343 SerLeuGluGlnValSerThrAlaLysSerGlnAspSerTrpArgThrAlaLeuSerThr 362
QY 1030 GCACACAGAGAAATAAATAAATAAATAAATAAATAAATAAATAA 1083
Db 363 SerGlnGluThrArgLysLeuValLysProLysAsnThrLysMetLysThrLys 380
RESULT 8
US-09-923-684-3
; Sequence 3, Application US/09923684
; Patent No. US20020081613A1
; GENERAL INFORMATION:
; APPLICANT: Narayanan, Ramaswamy
; TITLE OF INVENTION: ASSOCIATION OF SIM2 WITH CANCER
; FILE REFERENCE: 6818-24
; CURRENT APPLICATION NUMBER: US/09/923,684
; CURRENT FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-684-3
Alignment Scores:
Pred. No.: 7.6e-17 Length: 667
Score: 311.00 Matches: 124
Percent Similarity: 42.38% Conservative: 54
Best Local Similarity: 29.52% Mismatches: 123
Query Match: 15.47% Indels: 119
DB: 10 Gaps: 15
US-09-896-791B-2 (1-1100) x US-09-923-684-3 (1-667)
QY 64 AAGGAGAAGTCGGGGAGCGGGCCGAGCGGGCGGCGAGCAGGAGGAGGTGCTGTAC 123
Db 2 LysGluLysSerLysAsnAlaLysThrArgArgGluLysGluAsnGlyGluPheTyr 21
QY 124 CAGCTGGCGCACACTCTGCCCTTTGGCGCGGCTCAGCGCGCACCTGGACAGGCTCC 183
Db 22 GluLeuAlaLysLeuLeuProLeuProSerAlaIleThrSerGlnLeuAspLysAlaSer 41
QY 184 ATCATGCGCCTACAATCAGCTACTCGCATGCACCGCTCTCGCAGCAGGT----- 237
Db 42 IleIleArgLeuThrThrSerTyrLeuLysMetArgAlaValPheProGluGlyLeuGly 61
QY 238 -----GGAAAAGGGGAGAGCC-----ACTGGA 261
Db 62 AspAlaTrpGlyGlnProSerArgAlaGlyProLeuAspGlyValAlaLysGluLeuGly 81
QY 262 CGCTGCTACTGAAGCCCTGGAGGTTCGTATGTTACTACCCGCCGAGGAGACAT 321
Db 82 SerHisLeu-LeuGlnThrLeuAspGlyPheValPheValAlaSerAspGlyLysII 101
QY 322 GGCTTACTGTGGAAATGTCAGCAAGCACCTGGGCGCTCAGTCAGTGGACCTCTGTTC 381
Db 101 eMetTyrIleSerGluThrAlaSerValHisLeuGlyLeuSerGln----- 116
QY 382 TCCTCCCTGATACATAACCCCACTCTCTGGTACCAATTCCTCTGGAGCTCATTTGACAC 441
::: ||||| |||||
```

```
Db 117 -----Val--GluLeuThrGlyAsn 122
QY 442 AGTATCTTTTATCCATCTCTGTGACCAAGAGGAACCTTCAAGACGCCCTGACCCCC 501
Db 123 SerIleTyrGluTyrIleHisProSerAspHisAspGluMetThrAlaValLeuThrAla 142
QY 502 AGCCCGAACCTGTCAAAGAAGACTG---GAAGCCCCCAACAGAGCCCACTTTTCCCTG 558
Db 143 HisGlnProLeuHisHisLeuLeuGlnGluTyrGluIleGluArgSerPhePheLeu 162
QY 559 CGAATGAAGAGCAGCTCACACAGAGGCGGCACGCTCAACCTCAAGAGCGGCCACCTGG 618
Db 163 ArgMetLysCysValLeuAlaLys-----ArgAsnAlaGlyLeuThrCysSerGlyTyr 180
QY 619 AAGTGTCTGACCTGCTCAGGACATATGAGGCGCTACAAAGCCCCCTGCACAGACTCCCT 678
Db 181 LysValIleHisCysSerGlyTyrLeuLysIleArgGlnTyrMetLeuAspMetSerLeu 200
QY 679 GCCGGGAGC-----CCTCGCTCCGAGCCTCCCTG 708
Db 201 TyrAspSerCysTyrGlnIleValGlyLeuValAlaValGlyGlnSerLeuProPro--- 219
QY 709 CAATGCTGTGCTTATCTGTGAAGCATCCCGAGCTCCCTTCCACGATGGT----- 762
Db 220 -----SerAlaIleThrGluIleLysLeuTyrSerAsnMetPhe 232
QY 763 -----GCTACTCTGGGTCTT----- 777
Db 233 MetPheArgAlaSerLeuAspLeuLysLeuIlePheLeuAspSerArgValThrGluVal 252
QY 778 -----CCACAG-----GAGAAGACTCCCATC-----TCT 801
Db 253 ThrGlyTyrGluProGlnAspLeuIleGluLysThrLeuTyrHisHisValHisGlyCys 272
QY 802 ACCTTATTCACCTCTTTTGAAGGACACTACTTGTCTGCTCAAGAGTGGCTGT--- 858
Db 273 AspValPheHisLeuArgTyrAlaHisHisLeuLeuValLysGlyGlnValThrThr 292
QY 859 -----CAGTGTCTACAGGGGAAAGG----- 879
Db 293 LysTyrTyrArgLeuLeuSerLysArgGlyGlyTrpValTrpValGlnSerTyrAlaThr 312
QY 880 -----ACTGAATCTCTCTCTCC-----TCATGGGTCTTG 909
Db 313 ValValHisAsnSerArgSerArgProHisCysIleValSerValAsnTyrValLeu 332
QY 910 TGGGCCCTTAACCGGAAATAATTCCTGGCTAGGAGGAGTGAAGGACATGGCCACGTA 969
Db 333 ThrGluIleGluTyrLysGlu-----LeuGlnLeu 342
QY 970 TCCTTACCCAGAAACCCCAAAATGCTCCAAAACCAACCAATAAAGACCTCTCTTGTGTAG 1029
Db 343 SerLeuGluGlnValSerThrAlaLysSerGlnAspSerTrpArgThrAlaLeuSerThr 362
QY 1030 GCACACAGAGAAATAAATAAATAAATAAATAAATAAATAAATAA 1083
Db 363 SerGlnGluThrArgLysLeuValLysProLysAsnThrLysMetLysThrLys 380
RESULT 9
US-09-925-297-863
; Sequence 863, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
```

```

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (260)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-659

Alignment Scores:
Pred. No.: 1,21e-07 Length: 333
Score: 192.00 Matches: 40
Percent Similarity: 68.92% Conservative: 11
Best Local Similarity: 54.05% Mismatches: 19
Query Match: 9.55% Indels: 4
DB: 10 Gaps: 3

US-09-896-791B-2 (1-1100) x US-09-925-302-659 (1-333)
QY 538 ACAGAGCGCCACCTTCCCTGCGCAATGAAGACGACCTCACCAGCAGAGCGCGCAGCTC 597
Db 2 ThrGluArgAspPheMetArgMetLysCysThrValThrAsnArgGlyArgThrVal 21
QY 598 AACCTCAAGCGCCACCTGGAAGGTGCTGCACCTGCCTAGGACATATGAGGCGCTACAAG 657
Db 22 AsnLeuLysSerAlaThrTrpLysValLeuHisCysThrGlyGlnValLysValTyra 41
QY 658 CCC---CCTGCACAGACTCCCTCCGCGGGAGCCCTCGCTCGGAGCCTCCCTGCAATGC 714
Db 42 AsnCysProHisAsnSerLeuCysGly-----TyrLysGluProLeuLeuSerCys 59
QY 715 CTGGTGCTTATCTGTGAAGCCATCCCCCAGCTCCCTCCCTCCAC 756
Db 60 LeuIleMetCysGluProIle---GlnHisProSerHis 72
RESULT 11
US-09-764-864-1139
; Sequence 1139, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1139
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (338)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (376)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (465)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (485)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (491)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1139

Alignment Scores:
Pred. No.: 2,99e-05 Length: 591
Score: 162.00 Matches: 72
Percent Similarity: 39.43% Conservative: 53
Best Local Similarity: 22.71% Mismatches: 111
Query Match: 8.06% Indels: 81
DB: 10 Gaps: 3

```

```

1 APPLICANT: Penn, Sharron G.
2 APPLICANT: Rank, David R.
3 APPLICANT: Hanzel, David K.
4 APPLICANT: Chen, Wensheng
5 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID SEQUENCES
6 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
7 FILE REFERENCE: Aecomica-x-1
8 CURRENT APPLICATION NUMBER: US/09/864,761
9 CURRENT FILING DATE: 2001-05-23
10 PRIOR APPLICATION NUMBER: US 60/180,312
11 PRIOR FILING DATE: 2000-02-04
12 PRIOR APPLICATION NUMBER: US 60/207,456
13 PRIOR FILING DATE: 2000-05-26
14 PRIOR APPLICATION NUMBER: US 09/632,366
15 PRIOR FILING DATE: 2000-08-03
16 PRIOR APPLICATION NUMBER: GB 24263.6
17 PRIOR FILING DATE: 2000-10-04
18 PRIOR APPLICATION NUMBER: US 60/236,359
19 PRIOR FILING DATE: 2000-09-27
20 PRIOR APPLICATION NUMBER: PCT/US01/00666
21 PRIOR FILING DATE: 2001-01-30
22 PRIOR APPLICATION NUMBER: PCT/US01/00667
23 PRIOR FILING DATE: 2001-01-30
24 PRIOR APPLICATION NUMBER: PCT/US01/00664
25 PRIOR FILING DATE: 2001-01-30
26 PRIOR APPLICATION NUMBER: PCT/US01/00669
27 PRIOR FILING DATE: 2001-01-30
28 PRIOR APPLICATION NUMBER: PCT/US01/00665
29 PRIOR FILING DATE: 2001-01-30
30 PRIOR APPLICATION NUMBER: PCT/US01/00668
31 PRIOR FILING DATE: 2001-01-30
32 PRIOR APPLICATION NUMBER: PCT/US01/00663
33 PRIOR FILING DATE: 2001-01-30
34 PRIOR APPLICATION NUMBER: PCT/US01/00662
35 PRIOR FILING DATE: 2001-01-30
36 PRIOR APPLICATION NUMBER: PCT/US01/00661
37 PRIOR FILING DATE: 2001-01-30
38 PRIOR APPLICATION NUMBER: PCT/US01/00670
39 PRIOR FILING DATE: 2001-01-30
40 PRIOR APPLICATION NUMBER: US 60/234,687
41 PRIOR FILING DATE: 2000-09-21
42 PRIOR APPLICATION NUMBER: US 09/608,408
43 PRIOR FILING DATE: 2000-06-30
44 PRIOR APPLICATION NUMBER: US 09/774,203
45 PRIOR FILING DATE: 2001-01-29
46 NUMBER OF SEQ ID NOS: 49117
47 SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
48 SEQ ID NO 41589
49 LENGTH: 53
50 TYPE: PRT
51 ORGANISM: Homo sapiens
52 FEATURE:
53 OTHER INFORMATION: MAP TO AL139296.2
54 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3
55 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
56 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
57 OTHER INFORMATION: EXPRESSED IN HEPAT, SIGNAL = 2.7
58 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
59 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
60 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
61 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.9
62 OTHER INFORMATION: EST_HUMAN HIT: A1360153.1, EVALUAE 1.00e-21
63 OTHER INFORMATION: SWISSPROT HIT: Q24119, EVALUAE 5.00e-21
64 US-09-864-761-41589

```

Alignment Scores:	
Pred. No.:	8.73e-05
Score:	153.00
Percent Similarity:	76.00%
Best Local Similarity:	64.00%
Query Match:	7.61%
DB:	10
Length:	53
Matches:	32
Conservative:	6
Mismatches:	12
Indels:	0
Gaps:	0

RESULT 12
US-09-864-761-41589
; Sequence 41589, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 2, 2002, 20:32:35 : Search time 39 Seconds
(without alignments)
5422.967 Million cell updates/sec

Title: US-09-896-791B-2
Perfect score: 2010
Sequence: 1 gaattcgccagggccat.....aaaaaaaaaacatgcggccgc 1100

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_Spool/US09896791/runat_02122002_152658_5606/app_query.fasta_1.1287
-DB=PIR_73 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896791.ecgn_1_1_30_@runat_02122002_152658_5606 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THRPADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	982.5	48.9	667	JC7771	hypoxia inducible
2	610.5	30.4	811	JC4837	hypoxia-inducible
3	607.5	30.2	811	JC7619	hypoxia-inducible
4	606.5	30.2	826	J38972	hypoxia-inducible
5	597.5	29.7	813	JC5809	hypoxia-inducible
6	349	17.4	1505	JC4851	hypoxia-inducible
7	306.5	15.2	248	A58520	single-minded gene
8	242.5	12.1	655	A29945	neurogenesis regul
9	216.5	10.8	823	T21943	hypothetical prote
10	216.5	10.8	825	T21944	hypothetical prote
11	172.5	8.6	300	T24292	hypothetical prote
12	166.5	8.3	392	JC7633	aryl hydrocarbon n
13	162	8.1	805	JC7635	aryl hydrocarbon r
14	158.5	7.9	791	A56241	aryl hydrocarbon r

ALIGNMENTS

RESULT 1
JC7771
hypoxia inducible factor-3 alpha - human
C:Species: Homo sapiens (man)
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: JC7771
R:Hara, S.; Hamada, J.; Kobayashi, C.; Kondo, Y.; Imura, N.
Biochem. Biophys. Res. Commun. 287, 808-813, 2001
A:Title: Expression and characterization of hypoxia-inducible factor (HIF)-3alpha in
A:Reference number: JC7771; PMID:11573933
A:Contents: Kidney
A:Accession: JC7771
A:Molecule type: mRNA
A:Residues: 1-667 <HAR>
A:Cross-references: DDBJ:AB054067
C:Comment: This protein is a heterodimeric transcription factor that belongs to the b
lved in the regulation of hypoxia-inducible gene expression in human kidney.
C:Genetics:
A:Gene: hif-3alpha
A:Map position: 19
C:Keywords: kidney

Alignment Scores:	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1.38e-70	982.50	85.08%	82.66%	48.88%	2	US-09-896-791B-2 (1-1100) x JC7771 (1-667)	667	205	6	7	30	3
QY	43	AGGTGCAACACCGAGCTGCGGAGGAGAGTTCGGGACGCGCGCCCGGACGCGGCAGC	102									
Db	7	ArgSerThrThrGluLeuArgLysSerArgAspAlaAlaArgSerArgSer	26									
QY	103	CAGGAGACGGAGGTGCTGTACCGAGTGGCGGCACACTCTCCCTTTTTCGCGCGCGCGTCACG	162									

brain and muscle A
Ah receptor nuclea
aryl hydrocarbon r
Arnt-like PAS prot
aryl hydrocarbon r
aryl hydrocarbon r
aryl hydrocarbon r
aryl hydrocarbon r
aryl hydrocarbon r
aryl hydrocarbon r
aryl hydrocarbon r
cyclin T - fruit f
protein C4IG7.5 [1
hypothetical prote
brain and muscle A
tegument protein 6
neurogenesis regul
glucocorticoid rec
probable large pro
brain and muscle A
proline-rich prote
probable proline-r
unconventional myo
hypothetical prote
cell proliferation
hypothetical 70K p
ORF2 protein - Orf
eyelid - fruit fly

Db 27 GlnGluThrGluValLeuTyrgInLeuAlaHisThrLeuProPheAlaArgGlyValSer 46
QY 163 GCGCACCTGGCAAGGCTCATCATGCGCGCTCACAAATCAGCTACCTGCGCATGCCACCGC 222
|||||
Db 47 AlaHisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrrLeuArgMetHisArg 66
QY 223 CTCTGCGCAGCAGGT-----GGAAAGGGGGAGAGCCACTGGACGCCCTG 267
|||||
Db 67 LeuCysAlaAlaGlyGluTrpAsnGlnValGly-AlaGlyGlyGluProLeuAspAlaCys 86
QY 268 CTACCTGAAGGCCCTGGAGGTTTCGTATGCTACCTACCGCCGAGGGAGACATGCTTA 327
|||||
Db 86 sTyrrLeuLysAlaLeuGluGlyPheValMetValLeuThrAlaGluGlyAspMetAlaTy 106
QY 328 CTGTGCGAAATGTCAGCAAGCACCTGGGCCCTCAGTCAGTCAGTCGCTCTCTCTCTCCCTCC 387
|||||
Db 106 rLeuSerGluAsnValSerLysHisLeuGlyLeuSerGln----- 119
QY 388 CTGATACATAAACCCTCCTCGTACCAATTTCTCTCTGGAGCTCATTTGGACACAGTATC 447
|||||
Db 120 -----Leu--GluLeuIleGlyHisSerIle 127
QY 448 TTTGATTTTATCCATCCCTGTGACCAAGAGAACTTCAAGACGCCCTGACCCCGGCGG 507
|||||
Db 128 PheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThrProGlnGln 147
QY 508 AACCTGTCAAAGAAAGCTGGAAGCCCAACAGAGCGCCACTTTTCCTGCGCAATGAAG 567
|||||
Db 148 ThrLeuSerArgLysValGluAlaProThrGluArgCysPheSerLeuArgMetLys 167
QY 568 AGCAGGCTCACCAGCAGAGGGCGCAGCTCAACCTCAAAGCGGCCACCTGGAAGGTGCTG 627
|||||
Db 168 SerThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrpLysValLeu 187
QY 628 CACTGCTCAGGACATATAGGGGCTACAGCCCGCTGACAGACTTCCCTGCGCGGAGC 687
|||||
Db 188 AsnCysSerGlyHisMetArgAlaTyrrLysProProAlaGlnThrSerProAlaGlySer 207
QY 688 CTTGCTCCAGGCTCCCTGCAATGCTGCTGCTTATCTGTCGAAGCCATCCCGCAGCTC 747
|||||
Db 208 ProAspSerGluProLeuGlnCysLeuValIleCysGluAlaIlePro----- 225
QY 748 CCCTTCACGATGGTCT 765
|||||
Db 226 -----HisProGlySer 229
RESULT 2
JC4837
hypoxia-inducible factor 1 alpha - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jul-2000
C:Accession: JC4837
R:Wenger, R.H.; Rolfs, A.; Marti, H.H.; Guenet, J.L.; Gassmann, M.
Biochem. Biophys. Res. Commun. 223, 54-59, 1996
A:Title: Nucleotide sequence, chromosomal assignment and mRNA expression of mouse hypoxi
A:Reference number: JC4837; MUID:96254028; PMID:8660378
A:Accession: JC4837
A:Molecule type: mRNA
A:Residues: 1-810 <WEN>
A:Cross-references: EMBL:X95580; NID:g1430864; PIDN:CAA64833.1; PID:g4379202
A:Comment: This factor is involved in the oxygen-regulated transcription of several gene
C:Genetics:
A:Gene: Hif1alpha
A:Map position: 12
C:Keywords: transcription factor
F:5-58/Region: helix-loop-helix #status predicted
Alignment Scores:
Pred. No.: 7.84e-41 Length: 810
Score: 610.50 Matches: 129
Percent Similarity: 67.65% Conservative: 32
Best Local Similarity: 54.20% Mismatches: 49
Query Match: 30.37% Indels: 28

DB: 2 Gaps: 3
US-09-896-791b-2 (1-1100) x JC4837 (1-810)
QY 49 AACACCCAGCTCGGGAAGAGAGAGTCCGGGAGCGGCCGCGAGCGCGCAGCCAGGAG 108
|||||
Db 2 SerSerGluArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 21
QY 109 AGCGAGTCTGTACACAGTGGCGCACACTCTGCGCTTTTGGCGCGCGCTACAGCGCCAC 168
|||||
Db 22 SerGluValPheTyrrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 41
QY 169 CTGGACAAGCCCTCCATCATGCGCCCTCACAAATCAGCTACCTGCGCATGACCGCCCTGTC 228
|||||
Db 42 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrrLeuArgValArgLysLeuLeu 61
QY 229 GCAGCAGGTGGAAAAGGGGAGAG-----CCACTGGAGCCCTGCTACCT 273
|||||
Db 62 AspAlaGlyGly-LeuAspSerGluAspGluMetLysAlaGlnMetAspCysPheTyrrLe 81
QY 274 GAAGGCCCTGGAGGTTTCGTATGCTACTCACCGCGGAGGAGACATGGCTTACCTGTC 333
|||||
Db 81 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetValTyrrIleSe 101
QY 334 GCAAAATGTCAGCAAGCACCTGGCGCTCAGTCAGTCAGTCAGTCCTCTGTTCTCTCCCTCCCTGATA 393
|||||
Db 101 rAspAsnValAsnLysTyrrMetGlyLeuThrGln----- 112
QY 394 CATAACCCCACTCCTGGTACCAATTTCTCTGAGACTCATTTGGACACAGATATCTTTGAT 453
|||||
Db 113 -----Phe--GluLeuAlaGlyHisSerValPheAsp 122
QY 454 TTTATCCATCCCTGTGACCAAGAGAACTTCAAGACGCCCTGACCCCGAGCGGAACCTG 513
|||||
Db 123 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyPro 142
QY 514 TCAAGAAGAAGCTGGAAGCCCCAACAGACGCCACTTTTCCCTGCGAATGAAGACGACG 573
|||||
Db 143 ValArgLysGlyLysGluLeuAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 162
QY 574 CTCACACGAGCGGCGCAGCTCAACCTCAAGCGCGCCACCTGGAAGTGTGCTGCTGCTGCTG 633
|||||
Db 163 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 182
QY 634 TCAGGACATATGAGGCGCTACAAAGCCCGCTGCACAGACTTCCCGCTCGCGGAGCCCTCGC 693
|||||
Db 183 ThrGlyHisIleHisValTyrrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 200
QY 694 TCGAGCCCTCCCTGCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
|||||
Db 201 LysLysProProMetThrCysLeuValLeuIleCysGluProIlePro 216
RESULT 3
JC7619
hypoxia-inducible factor 1 alpha - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7619
R:Takahashi, T.; Sugishita, Y.; Nojiri, T.; Shimizu, T.; Yao, A.; Kinugawa, K.; Harad
Biochem. Biophys. Res. Commun. 281, 1057-1062, 2001
A:Title: Cloning of hypoxia-inducible factor 1 alpha cDNA from chick embryonic ventri
A:Reference number: JC7619; MUID:21134360; PMID:1123772
A:Contents: Embryonic ventricular myocytes
A:Accession: JC7619
A:Molecule type: mRNA
A:Residues: 1-811 <TAK>
A:Cross-references: DDBJ:AB013746
A:Comment: This factor belongs to the basic helix-loop-helix-Per/ARNT/Sim (HLH-PAS) f
eostasis of cells, and in redox stimuli.
C:Genetics:
A:Gene: hif-1alpha
C:Keywords: embryo; transcription factor
F:106-156/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PASI>

F: 249-299/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS2>
F: 762-811/Domain: conserved carboxy-terminal transactivation element #status predicted <PAS2>
F: 767-768/Region: conserved dilaucine repeat, important for oxygen-dependent degradation

Alignment Scores:		
Pred. No.:	1,36e-40	811
Score:	607.50	136
Percent Similarity:	67.35%	25
Best Local Similarity:	56.90%	Conservative: 51
Query Match:	30.22%	Mismatches: 28
DB:	2	Indels: 3
		Gaps: 3

US-09-896-791B-2 (1-1100) x JC7619 (1-811)

QY	43	AGTCTGACACACCAGCTGCGGAGGAGAGTAGTCGCGGGACGCGGCCCGCAGCCGGCGCAGC	102
		: : : :	
Db	12	ArgIleSerSerGluArgArgLysGluLysSerArgAspAlaAlaArgCysArgArgSer	31
		: : : :	
QY	103	CAGGACGAGGAGGTGCTGTACCAAGCTGGCGGCACACTCTCCCTTTTCGCGCGCGCTCAGC	162
		: : : :	
Db	32	LysGluSerGluValPheTyrGluLeuAlaHisGluLeuProLeuProHisThrValSer	51
		: : : :	
QY	163	GCACACTGGACAAAGCGCTCCATCATTCGCGCTCACAACTACGCTACCTCGCGCATGCACCGC	222
Db	52	AlaHisLeuAspLysAlaSerIleMetuArgLeuThrIleSerTyrLeuArgMetArgLys	71
		: : : :	
QY	223	CTCTGCGCAGCAGGTGGAAAGGGGGAGAG-----CCACTGGACGCGCTG	267

72	LeuLeuAspAlaGly-GluLeuGluThrGluAlaAsnMetGluLysGluLeuAsnCys	91
268	CTACTGAAGCCCTGAGGGTTTCGTATGGTACTACCGCGGAGGAGACATGCCTTA	327
91	eTyrLeuLysAlaLeuAspGlyPheValMetValLeuSerGluAspGlyAspMetIleTyr	111
328	CCGTGCGGAAATGTACGACAGACACCTGGGCCCTCACTAGT-GACCTCTGTCTCTCCTC	386
111	rMetSerGluAsnValAsnLysCysMetGlyLeuThrGlnPheAsp-	126

Qy 387 C C T G A T A C A T A A C C C C A C T C T C G T A C C A A T T T C T C T G G A C T C A T T G G A C A C A G A T A T 446
Db 127 -----LeuThrGLYHisSerVa 132

[illegible][illegible]

Db 172 sCysThrLeuThrSerArgGlyArgThrValAsnIleLysSerAlaThrTrpLysValle 192

Qy* 627 GCACCTGCTCAGGACACATATGAGGCGCTACAGACCCCTGGACACACTTCCCTGCGGGAG 686

[illegible]

RESULT 4
I38972
hypoxia-inducible factor 1 alpha - human

C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 05-Nov-1999
C:Accession: I38972; G01875
R:Wang, G.-L.; Jiang, B.-H.; Rue, E. A.; Semenza, G.L.
Proc. Natl. Acad. Sci. U.S.A. 92, 5510-5514, 1995
A:Title: Hypoxia-inducible factor 1 is a helix-loop-helix heterodimer
A:Reference number: I38972; PMID:95296340; PMID:7539918

A:Accession: I38972
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-826 <RES>
A:Cross-references: EMBL:U22431; NID:g881345; PIDN:AAC50152.1; PID:g881346
A>Note: parts of this sequence were confirmed by peptide sequencing
R:Hogenesch, J.B.; Chan, W.K.; Carver, L.A.; Bradford, C.A.
submitted to the EMBL Data Library, June 1995
A:Reference number: H00692
A:Accession: G01875
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-826 <HOG>
A:Cross-references: EMBL:U29165; NID:g1144012; PIDN:AAC51210.1; PID:g1144013
C:Genetics:
A:Gene: GDB:H1F1A
A:Cross-references: GDB:512229
A:Map position: 14q21-14q24
C:Keywords: heterodimer

Alignment Scores:	
Pred. No.:	1.64e-40
Score:	606.50
Percent Similarity:	68.07%
Best Local Similarity:	54.20%
Query Match:	30.17%
DB:	2
	3
	Gaps:
	Indels:
	28
	Mismatches:
	48
	Conservative:
	33
	Matches:
	129
	Length:
	826

US-09-896-791B-2 (1-1100) x I38972 (1-826)

[illegible]

Qy 169 CTGCACAGGCTCCATCATCGCGCTCACAAATCAGCTACTCTGGCATGCCACCGCCTTCG 238
|||||
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
|||||

QY	229	GCAGCAGGTGGAAAAAGGGGGAG	:			-----CCACTGGACGCGTGCACCT 277
Db	74	AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93				

Db	93	uLysAlaLeuAspGlyPheValMetValLeuThrAspGlyAspMetIleTyrIle	113
Qy	334	GGAAATGTGACGACGACCTGGGCTTCAGTCAGTGGACCTGTGTCCTGCTCCCTGATA	393

Db	113	raspAsnValAsnLysrMetGlyLeuThrGln-----	124
Qy	394	CATAAACCCCACTCCTGGTATCCAAATTTCTCTCTCGAGGCTCAATTTGGACACACATATCTTTGAT	453

DB	L25	-----Phe--GluLeuTrnGlyHissservAlpHeAsp
Qy	454	TTTATCATCCCTGTGCACCAAGAGGAACATTTCAGAGCGCCTGACCCCCAGCGCGAACCTG

Qy	514	TCAAAGAAGAGCTGGTAAGCCCCAACAGACAGCGGCACATTTCCCTGGCAATTAAGAGACAG	573
		:	
Db	155	ValLysLysGlyIysGluGlnAsnThrGlnArgSerPhePheIeuArgMetLysCysThr	174

[illegible]

Db 163 ArgMetLysCysValLeuAlaLys-----ArgAsnAlaGlyLeuThrCysSerGlyTyr 180

QY 619 AAGGTGCTGCACCTGCTCAGGACATATGAGG 648

Db 181 LysValIleHisCysSerGlyTyrLeuLys 190

RESULT 8

A29945

neurogenesis regulatory protein - fruit fly (Drosophila melanogaster) (fragment)

N:Alternate names: single-minded gene protein

C:Species: Drosophila melanogaster

C>Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 20-Mar-1998

C:Accession: A29945

Cell 52, 143-151, 1988

R: Crews, S.T.; Thomas, J.B.; Goodman, C.S.

A:Title: The Drosophila single-minded gene encodes a nuclear protein with sequence simil

A:Reference number: A29945; MUID:88151023; PMID:334560

A:Accession: A29945

A:Molecule type: mRNA

A:Residues: 1-655 <CRE>

A:Cross-references: GB:M19020; NID:gl58464; PID:gl58465

C:Genetics:

A:Gene: sim

A:Cross-references: FlyBase:FBgn0004666

C:Keywords: DNA binding; transcription regulation

Alignment Scores:

Pred. No.:	2,11e-11	Length:	555
Score:	242.50	Matches:	76
Percent Similarity:	45.56%	Conservative:	42
Best Local Similarity:	29.34%	Mismatches:	67
Query Match:	12.06%	Indels:	74
DB:	2	Gaps:	7

US-09-896-791B-2 (1-1100) x A29945 (1-655)

QY 124 CAGCTGGCGCACACTGTGCTTGGCGGCGTGCAGCGGCACCTGGACAGGCGCTCC 183

Db 4 GluLeuAlaLysLeuLeuProAlaIleThrSerGlnLeuAspLysAlaSer 23

QY 184 ATCATGGCGCTCACATACGTACCTGCGCATGCGCGCTCTGGCGCAGCAGGT----- 237

Db 24 ValIleArgLeuThrThrSerTyrLeuLysMetArgGlnValPheProAspGlyLeuGly 43

QY 238 -----GGAAAAAGGGGAGGACCCACTGGACCC 264

Db 44 GluAlaTrpGlySerSerProAlaMetGlnArgGlyAlaThrIleLysGluLeuGlySer 63

QY 265 CTGCTACCTGAAGGCCCTGGAGGTTTGTCTGCTGCTACTCAACCGCGGAGGAGACATGCG 324

Db 64 HisLeu-LeuGlnThrLeuAspGlyPheIlePheValValAlaProAspGlyLysIleMe 83

QY 325 TTACCTGTCGGAATGTTCAGCAGCAGCCTGGCGCTCAGTCACTAGTGGACCTCTGTTCCCTCC 384

Db 83 tTyrIleSerGluThrAlaSerValHisLeuGlyLeuSerGln----- 97

QY 385 TCCTGATACATAACCCCACTCTGTGTACCAATTTCTCTCTGAGCTCATTTGACACAGT 444

Db 98 -----Val--GluLeuThrGlyAsnSer 104

QY 445 ATCTTTGATTTTATCCATCCCTGTGACCAAGGAAGTTCAGAGCGCCCTGACCCCCAGG 504

Db 105 IlePheGluTyrIleHisAsnTyrAspGlnAspGluMetAsnAlaIleLeuSerLeuHis 124

QY 505 CGGAACCTCTCAAGAGAGAGCTG----- 528

Db 125 ProHisIleAsnGlnHisProLeuAlaGlnThrHisThrProIleGlySerProAsnGly 144

QY 529 ---GAAGCCCAACA-----GAGCCG 546

Db 145 ValGlnHisProSerAlaTyrAspHisAspArgGlySerHisThrIleGluLeuGly 164

QY 547 CACTTTTCCCTCGGAATGAAGAGCAGCTCACAGCAGAGGCGCAGCTCAACCTCAA 606

Db 165 ThrPhePheLeuArgMetLysCysValLeuAlaLys-----ArgAsnAlaGlyLeuThr 182

QY 607 GCGGCCACCTGGAAGGTGCTGCACCTGCCTCAGGACATATGAGGCGCTTACAGCCCTGCA 666

Db 183 ThrSerGlyPheLysValIleHisCysSerGlyTyrLeuLysAlaArgIleTyrProAsp 202

QY 667 CAGACTTCCCTGCGGAGGCGCTCGCTCCGAGCGCTCCCAATGCTGCTGCTTATC 726

Db 203 ArgGlyAspGlyGlnGlySer-----LeuIleGlnAsnLeuLeuVal 217

QY 727 TGT-----GAAGCATCCCTCCAGCTCCCTCCAC 756

Db 218 AlaValGlyHisSerLeuProSerAlaIleThrGluIleLysLeuHis 234

RESULT 9

T21943

hypothetical protein F38A6.3a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T21943; T26899

R: Mortimore, B.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19491

A:Accession: T21943

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-823 <WIL>

A:Cross-references: EMBL:Z92833; PIDN:CAB07380.1; GSPDB:GN00023; CESP:F38A6.3a

A:Experimental source: clone F38A6

R: Ainscough, R.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z20282

A:Accession: T26899

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-823 <WI2>

A:Cross-references: EMBL:AL023842; PIDN:CAA19520.1; GSPDB:GN00023; CESP:F38A6.3a

A:Experimental source: clone Y44A6D

C:Genetics:

A:Gene: CESP:F38A6.3a

A:Map position: 5

A:Introns: 19/2; 75/1; 111/2; 175/1; 226/3; 331/3; 518/1; 627/2; 711/1; 768/3

Alignment Scores:

Pred. No.:	2,55e-09	Length:	823
Score:	216.50	Matches:	69
Percent Similarity:	47.33%	Conservative:	46
Best Local Similarity:	28.40%	Mismatches:	66
Query Match:	10.77%	Indels:	63
DB:	2	Gaps:	9

US-09-896-791B-2 (1-1100) x T21943 (1-823)

QY 43 AGGTGGAACACCGAGTCTCGGAAGGAGAGTCTCGGAGCGCGCGCGCGCGCGCGCAGC 102

Db 110 LysArgAsnMetGluArgArgGluThrSerArgHisAlaAlaArgAspArgSer 129

QY 103 CAGGAGACGAGGTGCTGTACAGTGGGCGACACTCTGCGCTTTGCG---CGCGCGCTC 159

Db 130 LysGluSerAspIlePheAspAspLeuLysMetCysValProIleValGluGluGlyThr 149

QY 160 ACAGCGCACCTGGACAGGCTCCATCATCGCCCTCACATCAGTACCTACCTGCGCATGCAC 219

Db 150 ValThrHisLeuAspArgIleAlaLeuLeuArgValAlaAlaThrIleCysArgLeuArg 169

QY 220 CGCCTCTCGCGACGAGTGTGAAAAAGGGGA----- 250

Db 170 Lys-ThrAlaGlyAsnValLeuGluAsnAsnLeuAspAsnGluIleThrAsnGluValTr 189

QY 251 -----GAGCCACTGGAGCGCTGTACTGAGGCGCTGGAGGGTTTCGTGATGCTACT 303

Db 189 pThrGluAspThrIleAlaGluCys-----LeuAspGlyPheValMetIleVa 205

Qy	304	CACGCCGAGGGAGCATGCCTTACTGTTCGGAAAAATGTCAGCAAGCACCCTCGGCCCTCAG	363
Df	:	: : : : :	
Db	205	lAspSerAspSerSeriLeuTyrrValThrGluservAlalaMetTyrrLeuGlyLeuth	225
Qy	364	TcagTG - GAcCTCTGTTCCTCCCTGATAcATAccCcAcTCtCGGTaccAAntTTCTC	422
Df	:	:	
Db	225	rGlnThrAsp-----	228
Qy	423	TCTGGAGCTCATTTGGACACAGTATCTTTGATTTCCTTCCTGTGACACAAGAGAacT	482
Df	:	: : : : : :	
Db	229	- ---- LeuThrGlyArgAlaleuArgaspPheLeuHisProSerAspTyrrAspcLuPh	246
Qy	483	TCAAGAGCCCTCACCCCCAGCCGACCTGTCAAAGAAGAGCTGGAGAGCCCCA----	537
Df	:	: : :	
Db	246	e-----AspLysGlnserLysMetLeuHisLysProArggl	258
Qy	538	-----ACAGAGCGCCACTTTTCCTCGGAANTGAAGAGCACGCTCACcAG	581
Df	:	: : : : : :	
Db	258	yGluAspThrAspThrThrGlyleasnMetvalLeuArgMetLysThrvalIleserPr	278
Qy	582	CAGAGCGCGCAGCCTCAACCTCAAGCGCGCACCTGGAAG-----GTGCTGCA	629
Df	:	:	
Db	278	oAtgGlyArgCysLeuasnLeulyssSerAlaleutyrrLysServalSerPheLeuVaHi	298
Qy	630	C-----TGCTCAGGACATYTAGGGCGCTCAAGCGCCCCTGACAGACTTCccc	677
Df	:	:	
Db	298	sSerLysValserThrGlyglyHisValSerPheMetGin-----GlyileThrfilePr	316
Qy	678	TGCGCGG	684
Df	:	:	
Db	316	oAlaGly	318
RESULT 10			
T21944			
hypothetical protein F38A6.3b - Caenorhabditis elegans			
C:Species:Caenorhabditis elegans			
C>Date:15-Oct-1999 #sequence_revision15-Oct-1999 #text_change29-Oct-1999			
C:Accession:T21944; T26900			
R:Mortimore, B.			
submitted to the EMBL Data Library, March 1997			
A:Reference number: Z19491			
A:Accession:T21944			
A>Status: preliminary; translated from GB/EMBL/DDBJ			
A:Molecule type: DNA			
A:Residues: 1-825 <WIL>			
A:Cross-references: EMBL:P292833; PIDN:CAB07381.1; GSPDB:GN00023; CESP:F38A6.3b			
A:Xperimental source: clone F38A6			
R:Ainscough, R.			
submitted to the EMBL Data Library, June 1998			
A:Reference number: Z20282			
A:Accession:T26900			
A>Status: preliminary; translated from GB/EMBL/DDBJ			
A:Molecule type: DNA			
A:Residues: 1-825 <WT2>			
A:Cross-references: EMBL:AL023842; PIDN:CAA19521.1; GSPDB:GN00023; CESP:F38A6.3b			
A:Xperimental source: clone Y44A6D			
C:Genetics:			
A:Gene: CESP:F38A6.3b			
A:Map position: 5			
A:Introns: 19/2; 75/1; 111/2; 175/1; 226/3; 331/3; 518/1; 627/2; 713/1; 770/3			

Alignment Scores:	
Pred. No.:	2,55e-09
Score:	216.50
Percent Similarity:	47.33%
Best Local Similarity:	28.40%
Query Match:	10.77%
DB:	2
Length:	825
Matches:	69
Conservative:	46
Mismatches:	66
Indels:	63
Gaps:	9

US-09-896-791B-2 (1-1100) x T21944 (1-825)

QY 43 AGGTCGAACACCGAGCTGCGGAAGGAGAAAGTCGGGGACGGGGCCCGCAGCCGGCGCAGC 102

Seq. No.	Seq. Name	Seq. Type	Seq. Length	Seq. Description
110	LysArgAsnMetGluArgArgGluThrSerArgHisAlaAlaArgAspArgSer	Protein	129	hypotheical protein T01D3.2 - Caenorhabditis elegans
103	CAGGAGACGGAGGTGCTACACAGCTGGCGCACACTCTGCCCTTTGCG	Protein	159	C:Species: Caenorhabditis elegans
130	LysGluSerAspIlePheAspAspLeuLysMetCysValProIleValGluGluGlyThr	Protein	149	C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
160	AGCGCGCACCTGGACAGGCTCATCATCAGCGCTCACATCAGCTACCTGGCATGCAC	Protein	219	C:Accession: T24292
150	ValThrHisLeuAspArgIleAlaLeuLeuArgValAlaAlaThrIleCysArgLeuArg	Protein	169	R:Steward, C.
220	CGCTCTCCGACAGCTGGAAAGGGGA	Protein	250	submitted to the EMBL Data Library, October 1996
170	Lys-ThrAlaGlyAsnValLeuGluAsnAsnLeuAspAsnGluIleThrAsnGluValTr	Protein	189	A:Reference number: Z19870
251	GAGCCACTGGACGCTGCTACCTGAAGCGCTGGAGGTTTCGTATGTTACT	Protein	303	A:Accession: T24292
199	pThrGluAspThrIleAlaGluCys	Protein	205	A:Status: preliminary;
304	CACCGCGGAGGAGACATGCTTACCTTCGCGAAATGTGACGACACCTGGCGCTCAG	Protein	363	A:Molecule type: DNA
205	LaspSerAspSerIleLeuTyrValThrGluSerValAlaMetTyrLeuGlyLeuTh	Protein	225	A:Residues: 1-300 <WIL>
364	TCAGTG-GACCTCTGTTCTCTCCCTGATACATACCCCACTCTCTGGTACCAATTTCTC	Protein	422	A:Cross-references: EMBL:Z81110; PIDN:CA03258.1; GSPDB:GN00023; CESP:T01D3.2
225	rGlnThrAsp	Protein	228	A:Experimental source: clone T01D3
423	TCTGAGCTCATTCGACACAGTATCTTTGATTTTATCCATCCCTGTGACCAAGAGAACT	Protein	482	C:Genetics:
229	LeuThrGlyArgAlaLeuArgAspPheLeuHisProSerAspTyrAspGluPh	Protein	246	A:Gene: CESP:T01D3.2
483	TCAAGACGCCCTGACCCCGGCGGACCTGTCAAGAGAAGAGCTGGAGCCCA	Protein	537	A:Map position: 5
246	e	Protein	258	A:Introns: 69/3; 121/3
538	ACAGAGCCCACTTTCCTTCGCGAATGACGACACCTCACCAG	Protein	581	Alignment Scores:
258	yGluAspThrAspThrThrGlyIleAsnMetValLeuArgMetLysThrValIleSerPr	Protein	278	Pred. No.:
582	CAGAGGGCGCACCTCAACCTCAAGGGCGCACCTGGAAG	Protein	629	8.24e-06
278	oArgGlyArgCysLeuAsnLeuIleSerAlaLeuTyrLysSerValSerPheLeuValH	Protein	298	Length:
630	C	Protein	677	300
298	sSerLysValSerThrGlyGlyHisValSerPheMetGln	Protein	316	
678	TCGCGGG	Protein	684	
316	oAlaGly	Protein	318	
RESULT 11				
T24292				

Score: 172.50 Matches: 59
Percent Similarity: 46.82% Conservative: 44
Best Local Similarity: 26.82% Mismatches: 70
Query Match: 8.58% Indels: 47
DB: 2 Gaps: 6

US-09-896-791B-2 (1-1100) x F24292 (1-300)

QY 124 CAGCTGGCCACACTCTGGCGGCGGCTTGGCGGCGCTCAGCGCG---CACCTGGACAAGGCC 180
Db 8 GlnLeuAlaAsnGlnLeuProLeuAlaAlaAlaIleSerGlyGlnHisIleAspLysThr 27

QY 181 TCCATCATGCGCCTCACAATCAGCTACCTCGCATCGCATGCCCTCTGCGCAGCAGGTGGA 240
Db 28 ThrMetValArgLeuAlaThrAlaTyrIleLysLeuHisAsnIlePheGly-GlnSerGI 47

QY 241 AAAAGGGGAGACCACTGGACGCGCTCTAC----- 271
Db 47 nArgAlaTyrSerSerAlaAspTyrTyrTyrGlySerAspSerLeuTrpThrAsnAsnHi 67

QY 272 -CTGAAGCGCCTGGAGGCTTCGTCTACTCTACCTCAGCGCGGAGGAGACATGGCTTACCT 330
Db 67 sLeuAspLeuLeuAspGlyPhePheValIleLeuAspArgArgGlyAspValLeuTyrII 87

QY 331 GTCGGAAAATGTGACGAAGCACTGGCGCTCAGTCAGTCAGTGGACCTCTCTCTCTCCCTG 390
Db 87 eSerGluThrIleSerIleTyrLeuGlyLeuSerGln----- 99

QY 391 ATACATAACCCCACTCTCTGTGTACCAATTTCTCTCTGGAGCTCATTTGACACAGATATCTT 450
Db 100 -----Val--GluMetThrGlyAsnAlaMetVal 108

QY 451 GATTTTATCCATCCCTGTGACCAAGAGAACTTCAAGACGCGCCTGACCCCGCCGCGAAC 510
Db 109 AspTyrIleHisGluGlnAspIleAsnCysPheAsnSerAlaLeu----- 123

QY 511 CTGTCAAGAAGAGCTGGAGAGCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGAGC 570
Db 124 -----AsnTyrCysAspLeuAsnTrpProGlnMetCysAsnValArgValLysSer 140

QY 571 ACGCTCACCAGCAGAGCGCGCCTCAACTCAAGCGGCCACC---TGAAGGTGCTG 627
Db 141 SerLeuThrLysArgAlaAsnLysAspAlaValArgAlaSerProGlyTyrLysValLeu 160

QY 628 CACTGCTCAGGACATATGAGCGGCTCAACAGC-----CCCTCGCAGCAGACT 672
Db 161 ArgLeuGluIleThrMet-GlyProAsnThrAsnThrArgMetIleAlaCysTyrProMe 180

QY 673 TCCCTCGCGGAGCCCTCGCTCGGAGCCCTCCCTGCAATGCCCTGGTGCTTA 724
Db 180 tProThrProValLeuSerThrValThrIleProSerAsnSerPheValIle 197

RESULT 12
JC7633
aryl hydrocarbon nuclear translocator ARNT2-like factor, ARNT2X - zebra fish
C;Species: Brachydanio rerio (zebra fish)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C;Accession: JC7633
R;Hsu, H.J.; Wang, W.D.; Hu, C.H.
Biochem. Biophys. Res. Commun. 282, 487-492, 2001
A;Title: Ectopic expression of negative ARNT2 factor disrupts fish development.
A;Reference number: JC7633; PMID:21294759; PMID:11401485
A;Contents: Heart
A;Accession: JC7633
A;Molecule type: mRNA
A;Residues: 1-392 <HSU>
A;Cross-references: GB:AY007992
C;Comment: This factor, a basic helix-loop-helix PAS (bHLH-PAS) factor, heterodimerizes
C;Genetics:
A;Gene: arnt2X

Alignment Scores: 2.5e-05 Length: 392
Pred. No.:

Score: 166.50 Matches: 62
Percent Similarity: 40.43% Conservative: 50
Best Local Similarity: 22.38% Mismatches: 83
Query Match: 8.28% Indels: 83
DB: 2 Gaps: 9

US-09-896-791B-2 (1-1100) x JC7633 (1-392)

QY 64 AAGGAGAAGTCGCGGACGCGCGCCGCGCAGCGCGCAGCGAGGAGGTGCTGTAC 123
Db 47 ArgGluAsnHisSerGluIleGluArgArgArgAsnLysMetThrGlnTyrIleThr 66

QY 124 CAGCTGGCGCACACTCTGCCCTTTGGCGCGCGCTCAGCGCGCACCTGGACAAGCGCTCC 183
Db 67 GluLeuSerAspMetValProThrCysSerAlaLeuAlaArgLysProAspLysLeuThr 86

QY 184 ATCATGGCCTCACATCAGCTACCTACCTGCGCATCAGCGCTCTGCGCAGCAGGTGAAAA 243
Db 87 IleLeuArgMetAlaValSerHisMetLysSerMetArg-----Gly-Th 101

QY 244 AGGGGAGAGCCACTCGACGCGCTCTCTAC----- 271
Db 101 rGlyAsnThrSerThrAspGlyAlaTyrLysProSerPheLeuThrGluGlnGluLeu 121

QY 272 -----CTGAAGCGCCTGGAGGTTTCGTACTGTGTACTCACCCTCGAG---GGAGA 318
Db 121 sHisLeuIleLeuGluAlaAlaAspGlyPheLeuPheValValAlaGluThrGlyAr 141

QY 319 CATGGCTTACCTCTCGGAAAATCTCAGCAAGCACCTGGG---CCTCAGTCAGTGGACCT 374
Db 141 gValIleTyrValSerAspSerValThrProValLeuAsnHisProGlnSer----- 158

QY 375 CTGTTCTCTCTCCCTGTATACATAACCCCACTCTGCTGTACCAATTTCTCTCTGGAGCTCAT 434
Db 159 -----GluTrpPh 161

QY 435 TGGACACAGATCTTTGATTTTATCCATCCCTGTGACCAAGAGAACTTCAAGACGCGCT 494
Db 161 eGlySerThrLeuPheGluGlnValHisProAspValAspValAspLysLeuArgGluGlnLe 181

QY 495 GACCCCGAGCGCCCACTGTCAAGAAGAAGCTG----- 528
Db 181 uSerThrSerGluAsnSerMetThrGlyArgIleLeuAspLeuLysThrGlyThrVally 201

QY 528 ----- 528
Db 201 sLysGluGlyGlnGlnSerSerMetArgMetCysMetGlySerArgArgSerPheIleCy 221

QY 529 -----GAAGCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGAGCAC 572
Db 221 sArgMetArgCysGlySerAlaProLeuAsp---HisIleSerLeuAsnArgLeuSerSe 240

QY 573 GCTCACCAGCAGAGGCGGCACG-----CTCAACCTCAAGCGCCACCTGGAA 620
Db 240 rMetArgLysArgTyrArgAsnGlyLeuGlyProSerLysGluGlyGluAlaGlnTyrSe 260

QY 621 GGTGCTGCATCTGTCTCAGGACATATGAGGCGCTACAGCCCGCTCGACAGACTTCCCTC-- 678
Db 260 rValValHisCysThrGlyTyrIleLysAlaTrpProAlaGlyMetThrIleProAs 280

QY 679 -GCCGGAGCGCTCGCTCCGAGCGCTCCCTGCAATGCGCTGGCTGCTATC 726
Db 280 pGluAspThrGluAlaGlyGlnThrSerLysTyrCysLeuValAlaIle 296

RESULT 13
JC7635
aryl hydrocarbon receptor nuclear translocator 1 - chicken
N;Alternate names: hypoxia-inducible factor 1 beta
C;Species: Gallus gallus (chicken)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C;Accession: JC7635
R;Catron, T.; Mendiola, M.A.; Smith, S.M.; Born, J.; Walker, M.K.
Biochem. Biophys. Res. Commun. 282, 602-607, 2001

[illegible]

Search completed: December 2, 2002, 20:40:48
Job time : 49 secs

—

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 2, 2002, 19:48:55 ; Search time 23 Seconds

(without alignments)
3967.301 Million cell updates/sec

Title: US-09-896-791B-2

Perfect score: 2010

Sequence: 1 gaattcgccagcaggccat.....aaaaaaaaaacatcgccgc 1100

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US09896791/runat_02122002_152658_5574/app_query.fasta_1.1287
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=resp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896791.ecgn_1.1.13.@runat_02122002_152658_5574 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	613.5	30.5	822	1 HIFA_MOUSE	Q61221 mus musculus
2	606.5	30.2	826	1 HIFA_HUMAN	Q15665 homo sapien
3	593.5	29.5	874	1 PAS1_MOUSE	P97481 mus musculus
4	572.5	28.5	870	1 PAS1_HUMAN	Q99814 homo sapien
5	346	17.2	1507	1 SIM1_MOUSE	Q24167 drosophila
6	340.5	16.9	590	1 NPAL_MOUSE	Q99742 homo sapien
7	337	16.8	594	1 NPAL_MOUSE	P97459 mus musculus
8	324	16.1	657	1 SIM2_MOUSE	Q61079 mus musculus
9	312.5	15.5	766	1 SIM1_HUMAN	P81133 homo sapien
10	311	15.5	667	1 SIM2_HUMAN	Q14190 homo sapien
11	310.5	15.4	765	1 SIM1_MOUSE	Q61045 mus musculus
12	302.5	15.0	958	1 TRH_MOUSE	Q24119 drosophila
13	289.5	14.4	673	1 SIM_MOUSE	P05709 drosophila
14	165.5	8.2	712	1 ARN2_MOUSE	Q61324 mus musculus
15	165	8.2	706	1 ARN2_HUMAN	Q9Hb22 homo sapien
16	158.5	7.9	413	1 CYCL_MOUSE	Q61734 drosophila
17	158.5	7.9	791	1 ARNT_MOUSE	P53762 mus musculus
18	157.5	7.8	800	1 ARNT_RAT	P41739 rattus norv

19	156.5	7.8	789	1 ARNT_HUMAN	P27540 homo sapien
20	155.5	7.7	790	1 ARNT_RABIT	O02748 oryctolagus
21	153.5	7.6	583	1 BMAL_HUMAN	O00327 homo sapien
22	145	7.2	848	1 AHR_HUMAN	P35869 homo sapien
23	140	7.0	805	1 AHR_MOUSE	P30561 mus musculus
24	138.5	6.9	853	1 AHR_RAT	P41738 rattus norv
25	135.5	6.7	1097	1 CCT_MOUSE	O08433 drosophila
26	130.5	6.5	855	1 CLOC_MOUSE	O08785 mus musculus
27	129.5	6.4	644	1 ARNT_MOUSE	O15945 drosophila
28	128.5	6.4	1464	1 NCO2_HUMAN	O15596 homo sapien
29	126.5	6.3	846	1 CLOC_HUMAN	O15516 homo sapien
30	126	6.3	816	1 NPA2_MOUSE	P97460 mus musculus
31	126	6.3	937	1 CDA7_HUMAN	Q9un72 homo sapien
32	126	6.3	1462	1 NCO2_MOUSE	Q61026 mus musculus
33	125.5	6.2	1465	1 NCO2_RAT	Q9wu19 rattus norv
34	121.5	6.0	717	1 PRDD_HUMAN	Q9H4q3 homo sapien
35	121.5	6.0	1183	1 DRPL_RAT	P54258 rattus norv
36	121	6.0	824	1 NPA2_HUMAN	Q99743 homo sapien
37	120	6.0	1516	1 NCO2_XENLA	Q9W705 xenopus lae
38	119	5.9	950	1 CDA1_HUMAN	Q9V513 homo sapien
39	119	5.9	3530	1 MX15_HUMAN	Q9ukn7 homo sapien
40	118.5	5.9	215	1 Y215_ADE02	P03291 human adeno
41	117.5	5.8	649	1 V70K_EPMV	P20129 eggplant mo
42	117	5.8	948	1 CDA2_HUMAN	Q9Y5h9 homo sapien
43	116.5	5.8	640	1 ELI2_HUMAN	O00472 homo sapien
44	116.5	5.8	941	1 CDAC_HUMAN	Q9un75 homo sapien
45	114.5	5.7	443	1 GAT3_HUMAN	P23771 homo sapien

ALIGNMENTS

RESULT 1
HIFA_MOUSE
ID HIFA_MOUSE STANDARD; PRT: 822 AA.
AC Q61221; Q61665; Q61664; O08993; O08741;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (ARNT interacting protein).
DE HIF1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Hepatocytes;
RX MEDLINE=96355491; PubMed=8702901;
RA Li H., Ko H.P., Whitlock J.P. Jr.;
RT "Induction of phosphoglycerate kinase 1 gene expression by hypoxia.
RT Roles of Arnt and Hif1alpha".
RL J. Biol. Chem. 271:21262-21267(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=98034461; PubMed=9368100;
RA Luo G., Gu Y.-Z., Jain S., Chan K.M., Hogenesch J.B.,
RA Bradford C.A.;
RT "Molecular characterization of the murine Hif-1 alpha locus".
RL Gene Expr. 6:287-299(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=97354184; PubMed=9210478;
RA Wenger R.H., Rolfs A., Kvietikova I., Spielmann P., Zimmermann D.R.,
RA Gassmann M.;
RT "The mouse gene for hypoxia-inducible factor-1alpha. Genomic organization, expression and characterization of an alternative first exon and 5' flanking sequence".
RL Eur. J. Biochem. 246:155-165(1997).
RN [4]
RP SEQUENCE OF 13-822 FROM N.A.

QY 739 CCC 741
 Db 228 Pro 228
 RESULT 2
 HIFA_HUMAN STANDARD: PRT; 826 AA.
 AC Q16665;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (ARNT interacting protein) (Member of PAS protein 1) (MOP1) (HIF1 alpha).
 GS HIFA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 166-170; 259-289 AND 771-781.
 RX MEDLINE=95296340; PubMed=7539918;
 RA Wang G.L., Jiang B.H., Rue E.A., Semenza G.L.;
 RT "Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5510-5514(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hepatoma;
 RX MEDLINE=97236817; PubMed=9079689;
 RA Hogenesch J.B., Chan W.K., Jackiw V.H., Brown R.C., Gu Y.-Z.,
 RA Pray-Grant M., Perdev G.H., Bradford C.A.;
 RT "Characterization of a subset of the basic-helix-loop-helix-PAS superfamily that interacts with components of the dioxin signaling pathway".
 RL J. Biol. Chem. 272:8581-8593(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Rupert J.L., Hochachka P.W.;
 RL "Hif1a sequence in the Quechua, a high altitude population.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INVOLVED IN THE INDUCTION OF OXYGEN REGULATED GENES.
 CC SPECIFICALLY RECOGNIZES AN 8 BP HYPOXIA RESPONSE ELEMENT (HRE).
 CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER BHLH PROTEIN. HETERODIMER OF AN ALPHA AND A BETA (ARNT) SUBUNITS.
 CC THE ALPHA SUBUNIT IS UNIQUE TO HIF-1 WHEREAS HIF-1 BETA (ARNT) CAN DIMERIZE WITH OTHER BHLH-PAS PROTEINS. INTERACTS WITH HSP90.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES. HIGHEST LEVELS IN KIDNEY AND HEART.
 CC -!- INDUCTION: UNDER REDUCED OXYGEN TENSION.
 CC -!- DOMAIN: BOTH HYPOXIA RESPONSIVENESS AND TRANSACTIVATION CAPABILITY RESIDE WITHIN THE C-TERMINAL PART.
 CC -!- PTM: REQUIRES PHOSPHORYLATION FOR DNA-BINDING.
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS.
 CC -!- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U22431; AAC50152.1; -
 DR EMBL; U29165; AAC51210.1; -
 DR EMBL; AF207601; AAF20139.1; -
 DR EMBL; AF207602; AAF20140.1; -
 DR EMBL; AF208487; AAF20149.1; -
 DR TRANSFAC; T01610; -
 DR Genbank; HGNC:4910; HIF1A.

DR MIM: 603348; -
 DR InterPro: IPR001092; HLH_basic.
 DR InterPro: IPR001321; Hypoxindf1A.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000014; PAS_domain.
 DR Pfam: PF00785; PAC; 1.
 DR Pfam: PF00989; PAS; 2.
 DR PRINTS: PR01080; HYPOXIAIFIA.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS00888; HLH_2; 1.
 DR PROSITE; PS0112; PAS; 2.
 KW Repeat; DNA-binding; Nuclear protein; Transcription regulation;
 KW Activator; Phosphorylation.
 FT DNA_BIND 17 30 BASIC DOMAIN.
 FT DOMAIN 31 71 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 85 158 PAS 1.
 FT DOMAIN 228 298 PAS 2.
 FT DOMAIN 302 345 PAC.
 FT DOMAIN 615 621 POLY-THR.
 SQ SEQUENCE 826 AA; 92670 MW; ABD4F7DAA135BE2D CRC64;
 Alignment Scores:
 Pred. No.: 2,57e-35 Length: 826
 Score: 606.50 Matches: 129
 Percent Similarity: 68.07% Conservative: 33
 Best Local Similarity: 54.20% Mismatches: 48
 Query Match: 30.17% Indels: 28
 DB: 1 Gaps: 3
 US-09-896-791b-2 (1-1100) x HIFA_HUMAN (1-826)
 QY 49 AACACCGAGTGGCGAAGAGAAAGTGGCGGGACGGCGCGCGAGCGGCGAGCAGGAG 108
 Db 14 SerSerGluArgArgLysSerArgAspAlaAlaArgSerArgArgSerLysGlu 33
 QY 109 ACGGAGTGGCTACCGAGTGGCGGCACACTCTGCGCCTTTGCGCGGGCGTCCAGCGCGAC 168
 Db 34 SerGluValPheTyrGluLeuAlaHisGluLeuProLeuProHisAsnValSerHis 53
 QY 169 CTGGCAAGGCTCCATCGCTCAATCAGTCACTCGCATCGTCCAGTCCAGCGCTCTGTC 228
 Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
 QY 229 GCAGCAGTGGAAAGGGGGAGAG-----CCACTGGAGCGCTGCTACT 273
 Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrIle 93
 QY 274 GAAGCGCTGGAGGTTTCGTCATGTCTACCGCCGAGGAGGAGCATGGCTTACTGTC 333
 Db 93 uLysAlaLeuaspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIle 113
 QY 334 GAAATATGTACAGAACGACCTGGCGCTCAGTCAGTGGAGCTCTGTCTCTCTCTCTCTCTGATA 393
 Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
 QY 394 CATAACCCCATCTCTGGTACCATAATTTCTCTGAGAGTCTATGGAGTACACAGTATCTTGTAT 453
 Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
 QY 454 TTTATCCATCCTGTGACCAAGAGGAAGTTCAGAGCGCCCTCAGCCCGGAGCGGACCTG 513
 Db 135 PheThrHisProCysAspHisGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
 QY 514 TCAAAAGAGAGAGTGAAGACCCCAACAGAGCGCCACTTTTCTCCCTCGAATGAAGACGAG 573
 Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
 QY 574 CTCACAGCAGAGCGGCACGCTCAACCTCAAGCGGCCACCTGGAAGGTGCTGCACTGC 633
 Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194

Db	41	LeuProLeuProHisSerValSerSerHisLeuAspHisAlaSerIleMetArgLeuAla	60
Qy	199	ATCAGCTACCTGGCGATGCACCCGCTC-----TGGCCACAGGTGGAAAAGG	246
Db	61	IleSerPheLeuArgThrHisLysLeuLeuSerSerValCysSer-GluAsnGluSerG	80
Qy	247	GGGAGAG-----CCACTGGAGCGCTGCTACCTGAAGGCGCTGGAGGGTTTCGT	297
Db	80	uAlaGluAlaAspGlnMetAspAsnLeuTyLeuLysAlaLeuGluGluPheIleA	100
Qy	298	GGTACTCACCGCGAGGGAGACATGGCTTACCTGTCGGAATAATGTGACGACGAC	357
Db	100	aValValThrGlnAspGlyAspMetIlePheLeuSerGluAsnIleSerLysPheMet	120
Qy	358	CCTCAGTCACTGGACGCTCTGTTCTCTCCCTGATACATAAAGCCCACTCCTG	417
Db	120	yLeuThrGln-----	123
Qy	418	TTCTCTCTGGAGCTCATTTGGACACAGTACTTTGATTTTATCATCCCTGTGAC	477
Db	124	----Val-----GluLeuThrGlnHisSerIlePheAspPheThrHisProCys	141
Qy	478	GAAGTTCACAGACGCTGACCCCCAGG-----CCGAACCTGTCAAGAGAAGCT	531
Db	142	GluIleArgGluAsnLeuThrLeuLysAsnGlySerGlyPheGlyLysSerLysAsp	161
Qy	532	GCCCCACAGAGCGCCACTTTTCCCTCGCAATGAAGACGACGCTACACGACAG	591
Db	162	ValSerThrGluArgAspPhePheMetArgMetLysCysThrValThrAsnArgGly	181
Qy	592	ACGCTCAACCTCAAGGGCCACTGGAGGTGCTGCATGCTCAGGACATATGAGGCG	651
Db	182	ThrValAsnLeuLysSerAlaThrTPlsValLeuHisCysThrGlyGlnValArg	201
Qy	652	TACAAGGCC--CCTGCACAGACTCCCTCGCGGAGCCCTCGCTCCGAGGCTCC	708
Db	202	TyrAsnAsnCysProHisSerSerLeuCysGlySer-----LysGluProLeu	219
Qy	709	CAATGCTGGTGTCTATCTGTGAAGCATCCCCAGCTCCCTTCCAC	756
Db	220	SerCysLeuIleIleMetCysGluProIle--GlnHisProSerHis	234
RESULT 4			
ID	PAS1_HUMAN	STANDARD;	PRT; 870 AA.
AC	Q99814; Q99630;		
DC	15-DEC-1998 (Rel. 37, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DE	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Endothelial PAS domain protein 1 (EPAS1), a transcription factor		
DE	expressed in endothelial cells.*;		
GN	(MOP2).		
GN	EPAS1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=97152468; PubMed=9000051;		
RX	Tian H., McKnight S.L., Russell D.W.;		
RT	*Endothelial PAS domain protein 1 (EPAS1), a transcription factor		
RT	selectively expressed in endothelial cells.*;		
RL	Genes Dev. 11:72-82(1997).		
RN	[2]		
RN	SEQUENCE FROM N.A.		
RC	TISSUE=Hepatoma;		
RC	MEDLINE=97236817; PubMed=9079689;		
RA	Hogenesch J.B., Chan W.K., Jackiw V.H., Brown R.C., Gu Y.-Z.,		
RA	Pray-Grant M., Perdev G.H., Bradford C.A.;		
RT	*Characterization of a subset of the basic-helix-loop-helix-PAS		
RT	superfamily that interacts with components of the diolign signaling		
RT	pathway.*;		

```
US-09-896-791b-2 (1-1100) x PAS1_HUMAN (1-870)
QY 19 ATGGCGTGGGGTGCAGCGCGTGCAGTGCAGACACAGCGAGTGCAGGAGGAGAGTGCAGG 78
DB 1 MetThrAlaAspLysGluLysLysArgSerSerSerGluArgLysLysSerArg 20
QY 79 GAGCGCGCCGAGCGCGCGAGCCAGGAGGAGGAGTGCCTGTACCAAGTGCAGCGCACACT 138
DB 21 AspAlaAlaArgCysArgSerLysGluThrGluValPheTyrGluLeuAlaHisGlu 40
QY 139 CTGCGCTTTGCGCGCGTGCAGCGCGCAGCGCGCAGCGCGCTCCATCATGCGCCTCACA 198
DB 41 LeuProLeuProHisSerValSerHisLeuAspLysAlaSerIleMetArgLeuAla 60
QY 199 ATCAGCTACTGCGCATGCACCGCCCTC-----TGCGCAGCAGGTGGAAAAAGG 246
DB 61 IleSerPheLeuArgThrHisLysLeuLeuSerSerValCysSer-GluAsnGluSerG1 80
QY 247 GGGAGAG-----CCACTGGACCGCTGCTACCTGAAGCGCCCTGGAGGGTTTCGTGAT 297
DB 80 uAlaGluAlaAspGlnGlnMetAspAsnLeuTyrLeuLysAlaLeuGluGlyPheIleAl 100
QY 298 GGTACTCAGCGCGGAGAGACATGCTTACCTGTCGGAAATGCTCAGCAAGCAGCTGG 357
DB 100 aValAlThrGlnAspGlyAspMetIlePheLeuSerGluAsnIleSerLysPheMetG1 120
QY 358 CCTCAGTCACTGGACCTCTGTTCTCCTCCTCCTGATACATAACCCCACTCCTCGTACCAAT 417
DB 120 yLeuThrGln----- 123
QY 418 TTCTCTCTGGAGCTATTGGACACAGTATCTTTGATTTCATCCATCCCTGTGACCAAGAG 477
DB 124 ---Val--GluLeuThrGlyHisSerIlePheAspPheThrHisProcysAspHisGlu 141
QY 478 GAATTCAGACAGCCCTGACCCCCAGG-----CCGAACCTGTCAAGAGAGAGCTGGAA 531
DB 142 GluIleArgGluAsnLeuSerLeuLysAsnGlySerGlyPheGlyLysSerLysAsp 161
QY 532 GCCCCACAGAGCGCCACATTTTCCCTGCGAATGAAGCAGCAGCTCACCAGCAGAGGCGCG 591
DB 162 MetSerThrGluArgAspPheMetArgMetLysCysThrValThrAsnArgGlyArg 181
QY 592 ACCTCAACTCAAGCGGCCACTGGAAGGTGCTGCATGCTCAGGACATATGAGGGCC 651
DB 182 ThrValAsnLeuLysSerAlaThrTrpLysValLeuHisCysThrGlyGlnValLysVal 201
QY 652 TACAAGCCC---CCTGCACAGACTTCCCTGCGCGGAGCCCTCCGCTCCGAGCTCCCTCG 708
DB 202 TyrAsnAsnCysProHisAsnSerLeuCysGly-----TyrLysGluProLeuLeu 219
QY 709 CAATGCTGCTGTTATCTGTGAAGCATCCCGCAGCTCCCTCTCCAC 756
DB 220 SerCysLeuIleIleMetCysGluProIle---GlnHisProSerHis 234
RESULT 5
SIMA_DROME
ID SIMA_DROME STANDARD; PRT; 1507 AA.
AC Q24167; O9VAA5;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Similar protein.
GN SIMA OR CG7951.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96269413; PubMed=8682312;
RA Nambu J.R., Chen W., Hu S., Crews S.T.;
RT "The Drosophila melanogaster similar bHLH-PAS gene encodes a protein
```

```
RT related to human hypoxia-inducible factor 1 alpha and Drosophila
RT single-minded.";
RL Gene 172:249-254 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Bayraktaroglu C., Baldwin D.,
RA Ballesh R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
CC -!- FUNCTION: POSSIBLE DNA-BINDING TRANSCRIPTIONAL ACTIVATOR.
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED IN THE EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. HIGHEST TO HIF-1 ALPHA.
CC -!- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U43090; AAC47303.1; -.
DR EMBL; AE003772; AAF57008.2; -.
DR FlyBase; FBgn0015542; sima.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
```


US-09-896-791B-2 (1-1100) x NPAL_MOUSE (1-594)

Qy	61	CGGAAGGAGAAGTCGCGGGACGCGCGCGAGCGCGAGCAGAGAGCGAGGAGTCTGTG	120
Db			
Db	46	ArgLysGluLysSerArgAsnAlaAlaAArgTrpArgArgGlyLysGluAsnLeuGluPhe	65
Qy	121	TACCACTGTCGGCACACTCTGCGCCCTTTGCGCGCGCGGTGAGCGCGCACTGACAAAGGCC	180
Db			
Db	66	PheGluLeuAlaLysLeuLeuProLeuProGlyAlaIleSerSerGlnLeuAspLysAla	85
Qy	181	TCCATCATGCGCTCAACAATCACTACCTCGCATCCACCGCTCTGCCGACGAGT---	237
Db			
Db	86	SerIleValArgLeuSerValThrTyrLeuArgLeuArgPheAlaAlaLeuGlyAla	105
Qy	238	-----GGAAAAAGGGG	249
Db			
Db	106	ProProTrpGlyLeuArgAlaValGlyProProAlaGlyLeuAlaProGlyArgArgGly	125
Qy	250	-----AGGCCACTGGAGCGCTGACTACTGAAAGCC	279
Db			
Db	126	ProValAlaLeuValSerGluValPheGluGlnHisLeuGlyGlyHisIle-LeuGlnSe	145
Qy	280	CTGTGAGGGTTTGGTCATGTTACTACCGCCGAGGAGGACATGGCTTACTGTCGGAAAA	339
Db			
Db	145	rLeuAspGlyPheValPheAlaLeuAsnGlnGluGlyLysPheLeuTyrIleSerGluTh	165
Qy	340	TGTCAGCAAGCACCTGGCGCTCAGTCAGTGGACCTCTGTTCCTCCTCCGTATACATAAC	399
Db			
Db	165	rValSerIleTyrLeuGlyLeuSerGln-----	174
Qy	400	CCCACTCCTGGTACCAATTTCTCTGGAGCTCATTTGGACACAGTATCTTTGATTTATC	459
Db			
Db	175	-----Val-LeuLeuThrGlySerSerValPheAspTyrIle	186
Qy	460	CATCCCTGTCACCAAGAGGAACTTCAAGAGCCGCTG-----	495
Db			
Db	187	HisProGlyAspHisSerGluValLeuGluGlnLeuGlyLeuArgAlaAlaSerIleGly	206
Qy	496	-----ACCCCCAGCGCGCACTGTCAAAGAAG-----	522
Db			
Db	207	ProProThrPro--ProSerValSerSerSerSerSerSerSerSerSerLeuVal	225
Qy	523	-----AAGCTGGAAGCC-----CCAACA-----	543
Db			
Db	226	AspThrProGluIleGluAlaSerProThrGluAlaSerProAlaPheArgAlaGlnGlu	245
Qy	544	CGCCACTTTTCCCTGCGAATGAAGACGACGCTCACCAGAGAGGGCGCAGCTCAACCTC	603
Db			
Db	246	ArgSerPheValArgMetLysSerThrLeuThrLysArgGly-----LeuAsnVal	263
Qy	604	AAAGCGCCACCTGGGAAGTGTCTGCTCAGGACATATGAGGCGCTACAAG-----	657
Db			
Db	264	LysAlaSerGlyTyrLysValIleHisValThrGlyArgLeuArgAlaArgAlaLeuGly	283
Qy	658	-----CCCCCTGCACAGACTTCCCTCCTCCCGGAGCCCTCCGC	693
Db			
Db	284	LeuValAlaLeuGlyHisThrLeuProProAla-----ProLeu	296
Qy	594	TCCGAGCGCTCCCTGCAA-----	720
Db			
Db	297	AlaGluLeuProLeuHisGlyHisMetIleValPheArgLeuSerLeuGlyLeuThrIle	316
Qy	721	CTTATCTCTGAAGCC	735
Db			
Db	317	LeuAlaCysGluSer	321

RESULT 8
SIM2_MOUSE

ID	SIM2_MOUSE	STANDARD;	
AC	Q61079; Q61904; Q61046;	PRT;	657 AA.
DT	01-NOV-1997 (Rel. 35, created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		

DE Single-minded homolog 2 (SIM transcription factor) (mSIM).
OS Mus musculus (Mouse).
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=96299751; PubMed=8661115;
RA Moffett P., Dayo M., Reese M., McCormack M.K., Pelletier J.;
RT "Characterization of msim, a murine homologue of the Drosophila sim
transcription factor.";
RL Genomics 35:144-155(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=96144720; PubMed=8561800;
RA Ema M., Suzuki M., Morita M., Hirose K., Sogawa K., Matsuda Y.,
Gotoh O., Saijoh Y., Fujii H., Hamada H., Fujii-Kuriyama Y.;
RT "cDNA cloning of a murine homologue of Drosophila single-minded, its
mRNA expression in mouse development, and chromosome localization.";
RL Biochem. Biophys. Res. Commun. 218:588-594(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster;
RX MEDLINE=97020303; PubMed=8812055;
RA Fan C.M., Kuwana E., Bulfone A., Fletcher C.F., Copeland N.G.,
Jenkins N.A., Crews S., Martinez S., Puelles L., Rubenstein J.L.,
Tessier-Lavigne M.;
RT "Expression patterns of two murine homologs of Drosophila
single-minded suggest possible roles in embryonic patterning and in
the pathogenesis of Down syndrome.";
RL Mol. Cell. Neurosci. 7:1-16(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR X Swiss Webster; TISSUE=Embryo;
RX MEDLINE=96299750; PubMed=8661114;
RA Yamaki A., Noda S., Kudo H., Shindoh N., Maeda H., Minoshima S.,
Kawasaki K., Shimizu Y., Shimizu N.;
RT "The mammalian single-minded (SIM) gene: mouse cDNA structure and
diencephalic expression indicate a candidate gene for Down
syndrome.";
RL Genomics 35:136-143(1996).
RN [5]
RP SUBUNITS.
RX MEDLINE=97172525; PubMed=9020169;
RA Probst M.R., Fan C.M., Tessier-Lavigne M., Hankinson O.;
RT "Two murine homologs of the Drosophila single-minded protein that
interact with the mouse aryl hydrocarbon receptor nuclear
translocator protein.";
RL J. Biol. Chem. 272:4451-4457(1997).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT MAY BE A MASTER GENE OF CNS
DEVELOPMENT IN COOPERATION WITH ARNT. IT MAY HAVE PLEIOTROPIC
EFFECTS IN THE TISSUES EXPRESSED DURING DEVELOPMENT.
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
BHLH PROTEIN. HETERODIMER OF SIM2 AND ARNT.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: TRANSCRIPTS WERE DETECTED IN HIGH LEVELS IN
KIDNEY FOLLOWED BY SKELETAL MUSCLE AND LUNG. LOW LEVELS WERE FOUND
IN TESTIS, BRAIN AND HEART. IN EARLY FETAL DEVELOPMENT IT IS FOUND
IN CNS, DEVELOPING KIDNEY, TONGUE EPITHELIUM AND CARTILAGE
PRIMORDIA.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS.
CC -!- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U42554; AAB19098.1; -
DR EMBL: D63333; BAA09700.1; -
DR EMBL: U40576; AAA91202.1; -
DR EMBL: D64135; BAA11013.1; -
DR MGD; MGI:98307; Sim2.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00888; HLH_2; 1.
DR PROSITE; PS0112; PAS; 2.
KW Developmental protein; Neurogenesis; Nuclear protein; Repeat;
KW Transcription regulation; DNA-binding.
FT DNA_BIND 1 13 BASIC DOMAIN
FT DOMAIN 14 54 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT PAS 1.
FT PAS 2.
FT PAS 3.
FT PAS 4.
FT PAS 5.
FT PAS 6.
FT PAS 7.
FT PAS 8.
FT PAS 9.
FT PAS 10.
FT PAS 11.
FT PAS 12.
FT PAS 13.
FT PAS 14.
FT PAS 15.
FT PAS 16.
FT PAS 17.
FT PAS 18.
FT PAS 19.
FT PAS 20.
FT PAS 21.
FT PAS 22.
FT PAS 23.
FT PAS 24.
FT PAS 25.
FT PAS 26.
FT PAS 27.
FT PAS 28.
FT PAS 29.
FT PAS 30.
FT PAS 31.
FT PAS 32.
FT PAS 33.
FT PAS 34.
FT PAS 35.
FT PAS 36.
FT PAS 37.
FT PAS 38.
FT PAS 39.
FT PAS 40.
FT PAS 41.
FT PAS 42.
FT PAS 43.
FT PAS 44.
FT PAS 45.
FT PAS 46.
FT PAS 47.
FT PAS 48.
FT PAS 49.
FT PAS 50.
FT PAS 51.
FT PAS 52.
FT PAS 53.
FT PAS 54.
FT PAS 55.
FT PAS 56.
FT PAS 57.
FT PAS 58.
FT PAS 59.
FT PAS 60.
FT PAS 61.
FT PAS 62.
FT PAS 63.
FT PAS 64.
FT PAS 65.
FT PAS 66.
FT PAS 67.
FT PAS 68.
FT PAS 69.
FT PAS 70.
FT PAS 71.
FT PAS 72.
FT PAS 73.
FT PAS 74.
FT PAS 75.
FT PAS 76.
FT PAS 77.
FT PAS 78.
FT PAS 79.
FT PAS 80.
FT PAS 81.
FT PAS 82.
FT PAS 83.
FT PAS 84.
FT PAS 85.
FT PAS 86.
FT PAS 87.
FT PAS 88.
FT PAS 89.
FT PAS 90.
FT PAS 91.
FT PAS 92.
FT PAS 93.
FT PAS 94.
FT PAS 95.
FT PAS 96.
FT PAS 97.
FT PAS 98.
FT PAS 99.
FT PAS 100.
FT PAS 101.
FT PAS 102.
FT PAS 103.
FT PAS 104.
FT PAS 105.
FT PAS 106.
FT PAS 107.
FT PAS 108.
FT PAS 109.
FT PAS 110.
FT PAS 111.
FT PAS 112.
FT PAS 113.
FT PAS 114.
FT PAS 115.
FT PAS 116.
FT PAS 117.
FT PAS 118.
FT PAS 119.
FT PAS 120.
FT PAS 121.
FT PAS 122.
FT PAS 123.
FT PAS 124.
FT PAS 125.
FT PAS 126.
FT PAS 127.
FT PAS 128.
FT PAS 129.
FT PAS 130.
FT PAS 131.
FT PAS 132.
FT PAS 133.
FT PAS 134.
FT PAS 135.
FT PAS 136.
FT PAS 137.
FT PAS 138.
FT PAS 139.
FT PAS 140.
FT PAS 141.
FT PAS 142.
FT PAS 143.
FT PAS 144.
FT PAS 145.
FT PAS 146.
FT PAS 147.
FT PAS 148.
FT PAS 149.
FT PAS 150.
FT PAS 151.
FT PAS 152.
FT PAS 153.
FT PAS 154.
FT PAS 155.
FT PAS 156.
FT PAS 157.
FT PAS 158.
FT PAS 159.
FT PAS 160.
FT PAS 161.
FT PAS 162.
FT PAS 163.
FT PAS 164.
FT PAS 165.
FT PAS 166.
FT PAS 167.
FT PAS 168.
FT PAS 169.
FT PAS 170.
FT PAS 171.
FT PAS 172.
FT PAS 173.
FT PAS 174.
FT PAS 175.
FT PAS 176.
FT PAS 177.
FT PAS 178.
FT PAS 179.
FT PAS 180.
FT PAS 181.
FT PAS 182.
FT PAS 183.
FT PAS 184.
FT PAS 185.
FT PAS 186.
FT PAS 187.
FT PAS 188.
FT PAS 189.
FT PAS 190.
FT PAS 191.
FT PAS 192.
FT PAS 193.
FT PAS 194.
FT PAS 195.
FT PAS 196.
FT PAS 197.
FT PAS 198.
FT PAS 199.
FT PAS 200.
FT PAS 201.
FT PAS 202.
FT PAS 203.
FT PAS 204.
FT PAS 205.
FT PAS 206.
FT PAS 207.
FT PAS 208.
FT PAS 209.
FT PAS 210.
FT PAS 211.
FT PAS 212.
FT PAS 213.
FT PAS 214.
FT PAS 215.
FT PAS 216.
FT PAS 217.
FT PAS 218.
FT PAS 219.
FT PAS 220.
FT PAS 221.
FT PAS 222.
FT PAS 223.
FT PAS 224.
FT PAS 225.
FT PAS 226.
FT PAS 227.
FT PAS 228.
FT PAS 229.
FT PAS 230.
FT PAS 231.
FT PAS 232.
FT PAS 233.
FT PAS 234.
FT PAS 235.
FT PAS 236.
FT PAS 237.
FT PAS 238.
FT PAS 239.
FT PAS 240.
FT PAS 241.
FT PAS 242.
FT PAS 243.
FT PAS 244.
FT PAS 245.
FT PAS 246.
FT PAS 247.
FT PAS 248.
FT PAS 249.
FT PAS 250.
FT PAS 251.
FT PAS 252.
FT PAS 253.
FT PAS 254.
FT PAS 255.
FT PAS 256.
FT PAS 257.
FT PAS 258.
FT PAS 259.
FT PAS 260.
FT PAS 261.
FT PAS 262.
FT PAS 263.
FT PAS 264.
FT PAS 265.
FT PAS 266.
FT PAS 267.
FT PAS 268.
FT PAS 269.
FT PAS 270.
FT PAS 271.
FT PAS 272.
FT PAS 273.
FT PAS 274.
FT PAS 275.
FT PAS 276.
FT PAS 277.
FT PAS 278.
FT PAS 279.
FT PAS 280.
FT PAS 281.
FT PAS 282.
FT PAS 283.
FT PAS 284.
FT PAS 285.
FT PAS 286.
FT PAS 287.
FT PAS 288.
FT PAS 289.
FT PAS 290.
FT PAS 291.
FT PAS 292.
FT PAS 293.
FT PAS 294.
FT PAS 295.
FT PAS 296.
FT PAS 297.
FT PAS 298.
FT PAS 299.
FT PAS 300.
FT PAS 301.
FT PAS 302.
FT PAS 303.
FT PAS 304.
FT PAS 305.
FT PAS 306.
FT PAS 307.
FT PAS 308.
FT PAS 309.
FT PAS 310.
FT PAS 311.
FT PAS 312.
FT PAS 313.
FT PAS 314.
FT PAS 315.
FT PAS 316.
FT PAS 317.
FT PAS 318.
FT PAS 319.
FT PAS 320.
FT PAS 321.
FT PAS 322.
FT PAS 323.
FT PAS 324.
FT PAS 325.
FT PAS 326.
FT PAS 327.
FT PAS 328.
FT PAS 329.
FT PAS 330.
FT PAS 331.
FT PAS 332.
FT PAS 333.
FT PAS 334.
FT PAS 335.
FT PAS 336.
FT PAS 337.
FT PAS 338.
FT PAS 339.
FT PAS 340.
FT PAS 341.
FT PAS 342.
FT PAS 343.
FT PAS 344.
FT PAS 345.
FT PAS 346.
FT PAS 347.
FT PAS 348.
FT PAS 349.
FT PAS 350.
FT PAS 351.
FT PAS 352.
FT PAS 353.
FT PAS 354.
FT PAS 355.
FT PAS 356.
FT PAS 357.
FT PAS 358.
FT PAS 359.
FT PAS 360.
FT PAS 361.
FT PAS 362.
FT PAS 363.
FT PAS 364.
FT PAS 365.
FT PAS 366.
FT PAS 367.
FT PAS 368.
FT PAS 369.
FT PAS 370.
FT PAS 371.
FT PAS 372.
FT PAS 373.
FT PAS 374.
FT PAS 375.
FT PAS 376.
FT PAS 377.
FT PAS 378.
FT PAS 379.
FT PAS 380.
FT PAS 381.
FT PAS 382.
FT PAS 383.
FT PAS 384.
FT PAS 385.
FT PAS 386.
FT PAS 387.
FT PAS 388.
FT PAS 389.
FT PAS 390.
FT PAS 391.
FT PAS 392.
FT PAS 393.
FT PAS 394.
FT PAS 395.
FT PAS 396.
FT PAS 397.
FT PAS 398.
FT PAS 399.
FT PAS 400.
FT PAS 401.
FT PAS 402.
FT PAS 403.
FT PAS 404.
FT PAS 405.
FT PAS 406.
FT PAS 407.
FT PAS 408.
FT PAS 409.
FT PAS 410.
FT PAS 411.
FT PAS 412.
FT PAS 413.
FT PAS 414.
FT PAS 415.
FT PAS 416.
FT PAS 417.
FT PAS 418.
FT PAS 419.
FT PAS 420.
FT PAS 421.
FT PAS 422.
FT PAS 423.
FT PAS 424.
FT PAS 425.
FT PAS 426.
FT PAS 427.
FT PAS 428.
FT PAS 429.
FT PAS 430.
FT PAS 431.
FT PAS 432.
FT PAS 433.
FT PAS 434.
FT PAS 435.
FT PAS 436.
FT PAS 437.
FT PAS 438.
FT PAS 439.
FT PAS 440.
FT PAS 441.
FT PAS 442.
FT PAS 443.
FT PAS 444.
FT PAS 445.
FT PAS 446.
FT PAS 447.
FT PAS 448.
FT PAS 449.
FT PAS 450.
FT PAS 451.
FT PAS 452.
FT PAS 453.
FT PAS 454.
FT PAS 455.
FT PAS 456.
FT PAS 457.
FT PAS 458.
FT PAS 459.
FT PAS 460.
FT PAS 461.
FT PAS 462.
FT PAS 463.
FT PAS 464.
FT PAS 465.
FT PAS 466.
FT PAS 467.
FT PAS 468.
FT PAS 469.
FT PAS 470.
FT PAS 471.
FT PAS 472.
FT PAS 473.
FT PAS 474.
FT PAS 475.
FT PAS 476.
FT PAS 477.
FT PAS 478.
FT PAS 479.
FT PAS 480.
FT PAS 481.
FT PAS 482.
FT PAS 483.
FT PAS 484.
FT PAS 485.
FT PAS 486.
FT PAS 487.
FT PAS 488.
FT PAS 489.
FT PAS 490.
FT PAS 491.
FT PAS 492.
FT PAS 493.
FT PAS 494.
FT PAS 495.
FT PAS 496.
FT PAS 497.
FT PAS 498.
FT PAS 499.
FT PAS 500.
FT PAS 501.
FT PAS 502.
FT PAS 503.
FT PAS 504.
FT PAS 505.
FT PAS 506.
FT PAS 507.
FT PAS 508.
FT PAS 509.
FT PAS 510.
FT PAS 511.
FT PAS 512.
FT PAS 513.
FT PAS 514.
FT PAS 515.
FT PAS 516.
FT PAS 517.
FT PAS 518.
FT PAS 519.
FT PAS 520.
FT PAS 521.
FT PAS 522.
FT PAS 523.
FT PAS 524.
FT PAS 525.
FT PAS 526.
FT PAS 527.
FT PAS 528.
FT PAS 529.
FT PAS 530.
FT PAS 531.
FT PAS 532.
FT PAS 533.
FT PAS 534.
FT PAS 535.
FT PAS 536.
FT PAS 537.
FT PAS 538.
FT PAS 539.
FT PAS 540.
FT PAS 541.
FT PAS 542.
FT PAS 543.
FT PAS 544.
FT PAS 545.
FT PAS 546.
FT PAS 547.
FT PAS 548.
FT PAS 549.
FT PAS 550.
FT PAS 551.
FT PAS 552.
FT PAS 553.
FT PAS 554.
FT PAS 555.
FT PAS 556.
FT PAS 557.
FT PAS 558.
FT PAS 559.
FT PAS 560.
FT PAS 561.
FT PAS 562.
FT PAS 563.
FT PAS 564.
FT PAS 565.
FT PAS 566.
FT PAS 567.
FT PAS 568.
FT PAS 569.
FT PAS 570.
FT PAS 571.
FT PAS 572.
FT PAS 573.
FT PAS 574.
FT PAS 575.
FT PAS 576.
FT PAS 577.
FT PAS 578.
FT PAS 579.
FT PAS 580.
FT PAS 581.
FT PAS 582.
FT PAS 583.
FT PAS 584.
FT PAS 585.
FT PAS 586.
FT PAS 587.
FT PAS 588.
FT PAS 589.
FT PAS 590.
FT PAS 591.
FT PAS 592.
FT PAS 593.
FT PAS 594.
FT PAS 595.
FT PAS 596.
FT PAS 597.
FT PAS 598.
FT PAS 599.
FT PAS 600.
FT PAS 601.
FT PAS 602.
FT PAS 603.
FT PAS 604.
FT PAS 605.
FT PAS 606.
FT PAS 607.
FT PAS 608.
FT PAS 609.
FT PAS 610.
FT PAS 611.
FT PAS 612.
FT PAS 613.
FT PAS 614.
FT PAS 615.
FT PAS 616.
FT PAS 617.
FT PAS 618.
FT PAS 619.
FT PAS 620.
FT PAS 621.
FT PAS 622.
FT PAS 623.
FT PAS 624.
FT PAS 625.
FT PAS 626.
FT PAS 627.
FT PAS 628.
FT PAS 629.
FT PAS 630.
FT PAS 631.
FT PAS 632.
FT PAS 633.
FT PAS 634.
FT PAS 635.
FT PAS 636.
FT PAS 637.
FT PAS 638.
FT PAS 639.
FT PAS 640.
FT PAS 641.
FT PAS 642.
FT PAS 643.
FT PAS 644.
FT PAS 645.
FT PAS 646.
FT PAS 647.
FT PAS 648.
FT PAS 649.
FT PAS 650.
FT PAS 651.
FT PAS 652.
FT PAS 653.
FT PAS 654.
FT PAS 655.
FT PAS 656.
FT PAS 657.
FT PAS 658.
FT PAS 659.
FT PAS 660.
FT PAS 661.
FT PAS 662.
FT PAS 663.
FT PAS 664.
FT PAS 665.
FT PAS 666.
FT PAS 667.
FT PAS 668.
FT PAS 669.
FT PAS 670.
FT PAS 671.
FT PAS 672.
FT PAS 673.
FT PAS 674.
FT PAS 675.
FT PAS 676.
FT PAS 677.
FT PAS 678.
FT PAS 679.
FT PAS 680.
FT PAS 681.
FT PAS 682.
FT PAS 683.
FT PAS 684.
FT PAS 685.
FT PAS 686.
FT PAS 687.
FT PAS 688.
FT PAS 689.
FT PAS 690.
FT PAS 691.
FT PAS 692.
FT PAS 693.
FT PAS 694.
FT PAS 695.
FT PAS 696.
FT PAS 697.
FT PAS 698.
FT PAS 699.
FT PAS 700.
FT PAS 701.
FT PAS 702.
FT PAS 703.
FT PAS 704.
FT PAS 705.
FT PAS 706.
FT PAS 707.
FT PAS 708.
FT PAS 709.
FT PAS 710.
FT PAS 711.
FT PAS 712.
FT PAS 713.
FT PAS 714.
FT PAS 715.
FT PAS 716.
FT PAS 717.
FT PAS 718.
FT PAS 719.
FT PAS 720.
FT PAS 721.
FT PAS 722.
FT PAS 723.
FT PAS 724.
FT PAS 725.
FT PAS 726.
FT PAS 727.
FT PAS 728.
FT PAS 729.
FT PAS 730.
FT PAS 731.
FT PAS 732.
FT PAS 733.
FT PAS 734.
FT PAS 735.
FT PAS 736.
FT PAS 737.
FT PAS 738.
FT PAS 739.
FT PAS 740.
FT PAS 741.
FT PAS 742.
FT PAS 743.
FT PAS 744.
FT PAS 745.
FT PAS 746.
FT PAS 747.
FT PAS 748.
FT PAS 749.
FT PAS 750.
FT PAS 751.
FT PAS 752.
FT PAS 753.
FT PAS 754.
FT PAS 755.
FT PAS 756.
FT PAS 757.
FT PAS 758.
FT PAS 759.
FT PAS 760.
FT PAS 761.
FT PAS 762.
FT PAS 763.
FT PAS 764.
FT PAS 765.
FT PAS 766.
FT PAS 767.
FT PAS 768.
FT PAS 769.
FT PAS 770.
FT PAS 771.
FT PAS 772.
FT PAS 773.
FT PAS 774.
FT PAS 775.
FT PAS 776.
FT PAS 777.
FT PAS 778.
FT PAS 779.
FT PAS 780.
FT PAS 781.
FT PAS 782.
FT PAS 783.
FT PAS 784.
FT PAS 785.
FT PAS 786.
FT PAS 787.
FT PAS 788.
FT PAS 789.
FT PAS 790.
FT PAS 791.
FT PAS 792.
FT PAS 793.
FT PAS 794.
FT PAS 795.
FT PAS 796.
FT PAS 797.
FT PAS 798.
FT PAS 799.
FT PAS 800.
FT PAS 801.
FT PAS 802.
FT PAS 803.
FT PAS 804.
FT PAS 805.
FT PAS 806.
FT PAS 807.
FT PAS 808.
FT PAS 809.
FT PAS 810.
FT PAS 811.
FT PAS 812.
FT PAS 813.
FT PAS 814.
FT PAS 815.
FT PAS 816.
FT PAS 817.
FT PAS 818.
FT PAS 819.
FT PAS 820.
FT PAS 821.
FT PAS 822.
FT PAS 823.
FT PAS 824.
FT PAS 825.
FT PAS 826.
FT PAS 827.
FT PAS 828.
FT PAS 829.
FT PAS 830.
FT PAS 831.
FT PAS 832.
FT PAS 833.
FT PAS 834.
FT PAS 835.
FT PAS 836.
FT PAS 837.
FT PAS 838.
FT PAS 839.
FT PAS 840.
FT PAS 841.
FT PAS 842.
FT PAS 843.
FT PAS 844.
FT PAS 845.
FT PAS 846.
FT PAS 847.
FT PAS 848.
FT PAS 849.
FT PAS 850.
FT PAS 851.
FT PAS 852.
FT PAS 853.
FT PAS 854.
FT PAS 855.
FT PAS 856.
FT PAS 857.
FT PAS 858.
FT PAS 859.
FT PAS 860.
FT PAS 861.
FT PAS 862.
FT PAS 863.
FT PAS 864.
FT PAS 865.
FT PAS 866.
FT PAS 867.
FT PAS 868.
FT PAS 869.
FT PAS 870.
FT PAS 871.
FT PAS 872.
FT PAS 873.
FT PAS 874.
FT PAS 875.
FT PAS 876.
FT PAS 877.
FT PAS 878.
FT PAS 879.
FT PAS 880.
FT PAS 881.
FT PAS 882.
FT PAS 883.
FT PAS 884.
FT PAS 885.
FT PAS 886.
FT PAS 887.
FT PAS 888.
FT PAS 889.
FT PAS 890.
FT PAS 891.
FT PAS 892.
FT PAS 893.
FT PAS 894.
FT PAS 895.
FT PAS 896.
FT PAS 897.
FT PAS 898.
FT PAS 899.
FT PAS 900.
FT PAS 901.
FT PAS 902.
FT PAS 903.
FT PAS 904.
FT PAS 905.
FT PAS 906.
FT PAS 907.
FT PAS 908.
FT PAS 909.
FT PAS 910.
FT PAS 911.
FT PAS 912.
FT PAS 913.
FT PAS 914.
FT PAS 915.
FT PAS 916.
FT PAS 917.
FT PAS 918.
FT PAS 919.
FT PAS 920.
FT PAS 921.
FT PAS 922.
FT PAS 923.
FT PAS 924.
FT PAS 925.
FT PAS 926.
FT PAS 927.
FT PAS 928.
FT PAS 929.
FT PAS 930.
FT PAS 931.
FT PAS 932.
FT PAS 933.
FT PAS 934.
FT PAS 935.
FT PAS 936.
FT PAS 937.
FT PAS 938.
FT PAS 939.
FT PAS 940.
FT PAS 941.
FT PAS 942.
FT PAS 943.
FT PAS 944.
FT PAS 945.
FT PAS 946.
FT PAS 947.
FT PAS 948.
FT PAS 949.
FT PAS 950.
FT PAS 951.
FT PAS 952.
FT PAS 953.
FT PAS 954.
FT PAS 955.
FT PAS 956.
FT PAS 957.
FT PAS 958.
FT PAS 959.
FT PAS 960.
FT PAS 961.
FT PAS 962.
FT PAS 963.
FT PAS 964.
FT PAS 965.
FT PAS 966.
FT PAS 967.
FT PAS 968.
FT PAS 969.
FT PAS 970.
FT PAS 971.
FT PAS 972.
FT PAS 973.
FT PAS 974.
FT PAS 975.
FT PAS 976.
FT PAS 977.
FT PAS 978.
FT PAS 979.
FT PAS 980.
FT PAS 981.
FT PAS 982.
FT PAS 983.
FT PAS 984.
FT PAS 985.
FT PAS 986.
FT PAS 987.
FT PAS 988.
FT PAS 989.
FT PAS 990.
FT PAS 991.
FT PAS 992.
FT PAS 993.
FT PAS 994.
FT PAS 995.
FT PAS 996.
FT PAS 997.
FT PAS 998.
FT PAS 999.
FT PAS 1000.
FT PAS 1001.
FT PAS 1002.
FT PAS 1003.
FT PAS 1004.
FT PAS 1005.
FT PAS 1006.
FT PAS 1007.
FT PAS 1008.
FT PAS 1009.
FT PAS 1010.
FT PAS 1011.
FT PAS 1012.
FT PAS 1013.
FT PAS 1014.
FT PAS 1015.
FT PAS 1016.
FT PAS 1017.
FT PAS 1018.
FT PAS 1019.
FT PAS 1020.
FT PAS 1021.
FT PAS 1022.
FT PAS 1023.
FT PAS 1024.
FT PAS 1025.
FT PAS 1026.
FT PAS 1027.
FT PAS 1028.
FT PAS 1029.
FT PAS 1030.
FT PAS 1031.
FT PAS 1032.
FT PAS 1033.
FT PAS 1034.
FT PAS 1035.
FT PAS 1036.
FT PAS 1037.
FT PAS 1038.
FT PAS 1039.
FT PAS 1040.
FT PAS 1041.
FT PAS 1042.
FT PAS 1043.
FT PAS 1044.
FT PAS 1045.
FT PAS 1046.
FT PAS 1047.
FT PAS 1048.
FT PAS 1049.
FT PAS 1050.
FT PAS 1051.
FT PAS 1052.
FT PAS 1053.
FT PAS 1054.
FT PAS 1055.
FT PAS 1056.
FT PAS 1057.
FT PAS 1058.
FT PAS 1059.
FT PAS 1060.
FT PAS 1061.
FT PAS 1062.
FT PAS 1063.
FT PAS 1064.
FT PAS 1065.
FT PAS 1066.
FT PAS 1067.
FT PAS 1068.
FT PAS 1069.
FT PAS 1070.
FT PAS 1071.
FT PAS 1072.
FT PAS 1073.
FT PAS 1074.
FT PAS 1075.
FT PAS 1076.
FT PAS 1077.
FT PAS 1078.
FT PAS 1079.
FT PAS 1080.
FT PAS 1081.
FT PAS 1082.
FT PAS 1083.
FT PAS 1084.
FT PAS 1085.
FT PAS 1086.
FT PAS 1087.
FT PAS 1088.
FT PAS 1089.
FT PAS 1090.
FT PAS 1091.
FT PAS 1092.
FT PAS 1093.
FT PAS 1094.
FT PAS 1095.
FT PAS 1096.
FT PAS 1097.
FT PAS 1098.
FT PAS 1099.
FT PAS 1100.
FT PAS 1101.
FT PAS 1102.
FT PAS 1103.
FT PAS 1104.
FT PAS 1105.
FT PAS 1106.
FT PAS 1107.
FT PAS 1108.
FT PAS 1109.
FT PAS 1110.
FT PAS 1111.
FT PAS 1112.
FT PAS 1113.
FT PAS 1114.
FT PAS 1115.
FT PAS 1116.
FT PAS 1117.
FT PAS 1118.
FT PAS 1119.
FT PAS 1120.
FT PAS 1121.
FT PAS 1122.
FT PAS 1123.
FT PAS 1124.
FT PAS 1125.
FT PAS 1126.
FT PAS 1127.
FT PAS 1128.
FT PAS 1129.
FT PAS 1130.
FT PAS 1131.
FT PAS 1132.
FT PAS 1133.
FT PAS 1134.
FT PAS 1135.
FT PAS 1136.
FT PAS 1137.
FT PAS 1138.
FT PAS 1139.
FT PAS 1140.
FT PAS 1141.
FT PAS 1142.
FT PAS 1143.
FT PAS 1144.
FT PAS 1145.
FT PAS 1146.
FT PAS 1147.
FT PAS 1148.
FT PAS 1149.
FT PAS 1150.
FT PAS 1151.
FT PAS 1152.
FT PAS 1153.
FT PAS 1154.
FT PAS 1155.<

Db 143 HisGlnProTyrHisSerHisPheValGlnGluTyrGluIleGluArgSerPhePheLeu 162

Oy 559 CGAATGAAGACGACCTCCACCGAGAGGGCGCGCTCAACCTCAAGAGCGCCACCTGG 618
||||| ||| ||| |||

Db 163 ArgMetLysCysValLeuAlaLys-----ArgAsnAlaGlyLeuThrCysGlyGlyTyr 180
||||| ||| ||| |||

Oy 619 AAGGTGCTGCACTGCTCAGGACATATGAGGCGCTACAGCCCTGCACAGACTTCCCTC 678
||||| ||| ||| |||

Db 181 LysValIleHisCysSerGlyTyrLeuLysIleArgGlnTyrSerLeuAspMetSerPro 200
||||| ||| ||| |||

RESULT 12

TRH_DROME STANDARD; PRT; 958 AA.

AC Q24119; Q24165; Q9W0Q7;
01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Trachealess protein.

GN TRH OR C66883.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=96136712; PubMed=8557198;

RA Wilk R., Weizman I., Shilo B.-Z.;

RT "trachealess encodes a bHLH-PAS protein that is an inducer of

RT tracheal cell fates in Drosophila.";

RL Genes Dev. 10:93-102(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=96136713; PubMed=8557189;

RA Isaac D.D., Andrew D.J.;

RT "Tubulogenesis in Drosophila: a requirement for the trachealess gene

RT product.";

RL Genes Dev. 10:103-117(1996).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananidis P.G., Scherer S.E., Richards S., Ashburner M., Henderson S.N.,

RA George R.A., Lewis S.E., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Sutton G.G., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Klumel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

CC -1- FUNCTION: TRANSCRIPTION FACTOR, MASTER REGULATOR OF TRACHEAL CELL

CC FATES IN THE EMBRYO, NECESSARY FOR THE DEVELOPMENT OF THE SALIVARY

CC GLAND DUCT AND THE POSTERIOR SPIRACLES. IT MAY INDUCE A GENERAL

CC FATE OF BRANCHED TUBULAR STRUCTURES OF EPITHELIAL ORIGIN. TCO/TRH

CC HETERODIMERS ARE INVOLVED IN THE CONTROL OF BREATHLESS EXPRESSION.

CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER

CC BHLH PROTEIN. HETERODIMER WITH TGO.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2 AND

CC 3; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: TRACHEA, SALIVARY GLAND DUCTS, POSTERIOR

CC SPIRACLES (FILZKOEPER PRIMORDIA) AND A SUBSET OF CELLS IN THE CNS.

CC -1- DEVELOPMENTAL STAGE: DURING EMBRYOGENESIS, FIRST DETECTED IN THE

CC TRACHEAL PLACODES AT STAGE 8, AND EXPRESSION CONTINUES THROUGHOUT

CC EMBRYONIC AND LARVAL DEVELOPMENT. IN THE DEVELOPING SALIVARY

CC GLAND, EXPRESSION IS OBSERVED IN THE ENTIRE GLAND AT STAGE 9 AND

CC BY STAGE 12, EXPRESSION IS CONFINED TO THE SALIVARY DUCTS.

CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF

CC TRANSCRIPTION FACTORS.

CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.

CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS

CC IN POSITIONS 22 TO 34.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U33427; AAA96257.1; ALT_INIT.

CC EMBL; U42699; AAA96754.1; ALT_FRAME.

CC EMBL; AE003468; AAF47386.1; -

CC FlyBase; FBgn0003749; trh.

CC InterPro; IPR001092; HLH_basic.

CC InterPro; IPR001610; PAC.

CC InterPro; IPR000014; PAS_domain.

CC Pfam; PF00989; PAS; 2.

CC SMART; SM00353; HLH; 1.

CC SMART; SM00086; PAC; 1.

CC SMART; SM00091; PAC; 2.

CC PROSITE; PS00038; HLH_1; 1.

CC PROSITE; PS00888; HLH_2; 1.

CC PROSITE; PS0112; PAS; 2.

CC Developal protein; Nuclear protein; Transcription regulation;

CC Repeat; DNA-binding; Alternative splicing.

CC DNA_BIND 86 99 BASIC DOMAIN.

CC DOMAIN 100 140 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).

CC DOMAIN 174 244 PAS 1.

CC DOMAIN 391 461 PAS 2.

CC DOMAIN 465 508 PAC.

CC DOMAIN 629 636 POLY-SER.

CC DOMAIN 154 157 POLY-GLY.

CC DOMAIN 251 254 POLY-SER.

CC DOMAIN 255 264 POLY-GLY.

CC DOMAIN 722 728 POLY-GLN.

CC VARSPLOC 281 286 MISSING (IN ISOFORM 2).

CC VARSPLOC 328 356 MISSING (IN ISOFORM 3).

CC CONFLICT 78 78 P -> A (IN REF. 1).

CC CONFLICT 250 250 G -> GG (IN REF. 2).

CC CONFLICT 703 703 A -> T (IN REF. 1).

FT CONFLICT 708 708 A -> P (IN REF. 1).
FT CONFLICT 829 829 A -> V (IN REF. 1).
SQ SEQUENCE 958 AA; 102241 MW; 8F9CF758F1370541 CRC64;

Alignment Scores:

Pred. No.: 5,68e-14 Length: 958
Score: 302.50 Matches: 96
Percent Similarity: 42.26% Conservative: 46
Best Local Similarity: 28.57% Mismatches: 72
Query Match: 15.05% Indels: 122
DB: 1 Gaps: 8

US-09-896-791b-2 (1-1100) x TRH_DROME (1-958)

QY 55 GAGCTCGGAAGAGAGTCTCGGGACGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 114
|||||
Db 85 GluLeuArgGlyGlySerArgAspAlaAlaArgSerArgArgGlyGlyGluAsnTyr 104
|||||
QY 115 GTGCTGTACCAAGTGGCGACACTCTGCGCTTTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGAC 174
|||||
Db 105 GluPheTyrGluLeuAlaLysMetLeuProLeuProAlaAlaIlePheSerGlnLeuAsp 124
|||||
QY 175 AGGCGCTCATCATGCGCTCACATCAGTACCTACCTGCGCGCATGCAC----- 219
|||||
Db 125 LysAlaSerIleIleArgLeuThrIleSerTyrLeuLysLeuArgAspPheSerGlyHis 144
|||||
QY 220 -----CGCTCTGCGCGCAGCGTGA 240
|||||
Db 145 GlyAspProTrpThrArgGluAlaSerSerSerSerLysLeuLysSerAlaAlaIle 164
|||||
QY 241 AAAAGGGGAGGCCACT-----GGACGCGCTGCTACCTGAAGGC 279
|||||
Db 165 ArgArgSerProAlaValAspLeuPheGluGlnHisGlnGlyThrHisIle-LeuGlnSe 184
|||||
QY 280 CTGGAGGCTTTCGTACTGCTACTCACCCTGCGGAGGAGACATGGCTACTCTGCGGAAA 339
|||||
Db 184 rLeuAspGlyPheAlaLeuAlaValAlaAlaAspGlyArgPheLeuTyrIleSerGluTh 204
|||||
QY 340 TGTACAGACGACCTGGGCGCTCAGTCACTGGAGGCGCTGTCTCTCTCTCTCTCTGATACATAAC 399
|||||
Db 204 rValSerIleTyrLeuGlyLeuSerGln----- 213
|||||
QY 400 CCGACTCTCGTACCAATTTCTCTGAGGAGCTATGGACACAGTATCTTGTATTTATC 459
|||||
Db 214 -----Val-GluMetThrGlySerSerIlePheAspTyrIle 225
|||||
QY 460 CATCTCTGTGACCAAGACACTTCAAGACGCCCTG----- 495
|||||
Db 226 HisGlnAlaAspHisSerGluIleAlaAspGlnLeuGlyLeuSerLeuThrSerGlyGly 245
|||||
QY 495 ----- 495
Db 245 GlyGlyGlyGlySerSerSerSerSerGlyGlyGlyGlyGlyGlyGlyGlyMet 265
|||||
QY 496 -----ACCCCGCGCGC 507
|||||
Db 266 AlaSerProThrSerGlyAlaSerAspGlySerGlyThrHisGlyThrAsnAsnPro 285
|||||
QY 508 AACCTGTCAAAGAGAGCTGGAAGCCCAACA-----GAGCGCCAC 549
|||||
Db 286 AspValAlaAlaSerMetThrGlnAlaSerThrSerGlyTyrLysGlyTyrAspArgSer 305
|||||
QY 550 TTTTCTCTCGAATGAAGACAGCTCACACAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
|||||
Db 306 PheCysValArgMetLysSerThrLeuThrLysArgGly-----CysHisPheLysSer 323
|||||
QY 610 GCCACTTGAAGATGGCTG-----CACTGCTCAGCA 639
|||||
Db 324 SerGlyTyrArgAlaSerAspAlaThrSerAsnCysAsnAsnGlyAsnAlaSerAsn 343
|||||
QY 640 CATATGAGGCGCTACAGCCCTCGCA----- 666
|||||
Db 344 AsnAlaLysAsnValLysAsnProGlySerAsnTyrSerValValLeuLeuLeuLys 363

QY 667 -----CAGACTTCCCTGCGGGAGCGCTCGTCCGAGCGCTCCCTGCAATGCCG 717
|||||
Db 364 LeuArgProGlnTyrThrPheSerHisSerArgLysSerGlnProLeuLeuGlyMet 383
|||||
QY 718 GTGCTTATCTGTGAGCCATCCCGAGCTCCCTTCCACGAT 759
|||||
Db 384 ValAlaLeuAlaIleAlaLeuProProProSerValHisGlu 397
|||||
RESULT 13
SIM_DROME
ID SIM_DROME STANDARD; PRT; 673 AA.
AC P05709; O96521; Q9VF23;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Single-minded protein.
GN SIM OR CG7771.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;
RN [1]
RP SEQUENCE OF 19-673 FROM N.A.
RX MEDLINE=88151023; PubMed=3345560;
RA Crews S.T., Thomas J.B., Goodman C.S.;
RT "The Drosophila single-minded gene encodes a nuclear protein with
sequence similarity to the per gene product.";
RL Cell 52:143-151(1988).
RN [2]
RP SEQUENCE OF 1-18 FROM N.A., AND SIMILARITY TO HLH PROTEINS.
RX MEDLINE=92103681; PubMed=1760843;
RA Nambu J.R., Lewis J.O., Wharton K.A. Jr., Crews S.T.;
RT "The Drosophila single-minded gene encodes a helix-loop-helix protein
that acts as a master regulator of CNS midline development.";
RL Cell 67:1157-1167(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99054545; PubMed=9840810;
RA Kasai Y., Stahl S., Crews S.;
RT "Specification of the Drosophila CNS midline cell lineage: direct
control of single-minded transcription by dorsal/ventral patterning
genes.";
RL Gene Expr. 7:171-189(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davidson L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

Search completed: December 2, 2002, 20:37:04
Job time : 35 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 2, 2002, 20:29:55 ; Search time 61 Seconds
(without alignments)
7431.207 Million cell updates/sec

Title: US-09-896-791B-2

Perfect score: 2010

Sequence: 1 gaattcgacagggccat.....aaaaaaaaaacatgcggcgc 1100

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p model -DEV=xlp
-Q=/cgn2_1/USPTO_SPOOL/US09896791/runat_02122002_152658_5586/app_query.fasta_1.1287
-DB=SPTREMBL_21 -OFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896791 -ACGN_1_1_66.@runat_02122002_152658_5586 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=20 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_21.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.virus.*
13: sp.vertibrate.*
14: sp.unclassified.*
15: sp.rvirus.*
16: sp.bacteriap.*
17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1636	81.4	307	11 Q8VHR1	Q8VHR1 mus musculu

2	1045	52.0	630	11	Q9QX54	Q9qxs4 mus musculu
3	1045	52.0	662	11	Q9Z2I5	Q9z2i5 mus musculu
4	1015.5	50.5	632	4	Q8WXA1	Q8wxal homo sapien
5	1011.5	50.3	632	4	Q96K34	Q96k34 homo sapien
6	1007	50.1	363	4	Q8TAP5	Q8tap5 homo sapien
7	1001	49.8	662	11	Q9JHS2	Q9jhs2 rattus norv
8	982.5	48.9	667	4	Q9Y2N7	Q9y2n7 homo sapien
9	753.5	37.5	648	4	Q9HAI2	Q9hai2 homo sapien
10	672	33.4	199	4	Q9HAM5	Q9ham5 homo sapien
11	613.5	30.5	258	11	Q9CYA8	Q9cya8 mus musculu
12	613.5	30.5	836	11	Q8R385	Q8r385 mus musculu
13	611.5	30.4	823	6	Q9XAT5	Q9xta5 bos taurus
14	607.5	30.2	811	13	Q9YIB9	Q9yib9 gallus gall
15	606.5	30.2	735	4	Q9GPT9	Q9gpt9 homo sapien
16	606.5	30.2	826	4	Q9UPB1	Q9upb1 homo sapien
17	603.5	30.0	823	11	Q9WTU9	Q9wtu9 rattus norv
18	593.5	29.5	825	11	Q35800	Q35800 rattus norv
19	587.5	29.2	874	11	Q9JHS1	Q9jhs1 rattus norv
20	575.5	28.6	867	13	Q9W7C6	Q9w7c6 gallus gall
21	575.5	28.6	870	13	Q9PTB3	Q9ptb3 coturnix co
22	572.5	28.5	870	6	Q9XTA4	Q9xta4 bos taurus
23	569.5	28.3	237	4	Q9H7Z9	Q9h7z9 homo sapien
24	565	28.1	873	13	Q8QGM4	Q8qgm4 fundulus he
25	547.5	27.2	766	13	Q98SW2	Q98sw2 oncorhynch
26	511.5	25.4	115	4	Q95262	Q95262 homo sapien
27	501.5	25.0	112	4	Q9UPH7	Q9uph7 homo sapien
28	424	21.1	805	13	Q918A9	Q918a9 xenopus lae
29	360	17.9	65	11	Q8R4D6	Q8r4d6 mus musculu
30	343.5	17.1	590	4	Q9BY83	Q9by83 homo sapien
31	327	16.3	925	11	Q9QZQ0	Q9qzq0 mus musculu
32	321	16.0	657	11	Q35391	Q35391 mus musculu
33	319	15.9	901	4	Q9H323	Q9h323 homo sapien
34	318	15.8	903	4	Q9BY81	Q9by81 homo sapien
35	311	15.5	902	5	Q8SX13	Q8sx13 drosophila
36	309.5	15.4	585	13	Q8UVV3	Q8uvv3 brachydanio
37	308.5	15.3	765	11	Q70284	Q70284 mus musculu
38	300.5	15.0	760	13	Q9DDU6	Q9ddu6 xenopus lae
39	294.5	14.7	745	13	Q98SJ5	Q98sj5 brachydanio
40	293.5	14.6	849	5	O15984	O15984 bombyx mori
41	269.5	13.4	103	6	Q9N110	Q9n110 ovis aries
42	258.5	12.9	108	11	Q9QZJ8	Q9qzj8 cavia porce
43	216.5	10.8	719	5	Q963J8	Q963j8 caenorhabdi
44	216.5	10.8	823	5	O45486	O45486 caenorhabdi
45	216.5	10.8	825	5	Q9TVM0	Q9tvm0 caenorhabdi

ALIGNMENTS

RESULT 1
Q8VHR1 PRELIMINARY; PRT; 307 AA.
ID Q8VHR1
AC Q8VHR1
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inhibitory PAS domain protein.
GN IPAS
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX PubMed=11734856;
RA Makino Y., Cao R., Svensson K., Bertilsson G., Asman M., Tanaka H.,
Cao Y., Berkenstam A., Poellinger L.;
RT "Inhibitory PAS domain protein is a negative regulator of hypoxia-
inducible gene expression.";
RL Nature 414:550-554(2001).
DR EMBL; AF416641; AAL39015.1; -.
DR InterPro; IPR001092; HLH_basic.
DR SMART; SM00353; HLH; 1.

DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 307 AA; 33990 MW; D89D3B225C9C3967 CRC64;

Alignment Scores:

Pred. No.: 3,61e-136 Length: 307
Score: 1636.00 Matches: 307
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 81.39% Indels: 0
DB: 11 Gaps: 0

US-09-896-791B-2 (1-1100) x Q8VHR1 (1-307)

QY 19 ATGCGTGGTGGCGTGCAGCGCGTGCAGTGAACACCGAGTGCAGGAGGAGAGTGCAGG 78
DB 1 MetAlaLeuGlyLeuGlnArgValArgSerGlnGluThrGluLeuArgLysGluLysSerArg 20
QY 79 GACGCGCGCGCGAGCGCGCGAGCCAGGAGGAGGAGTGTGTACCAGTGCAGCGCACACT 138
DB 21 AspAlaAlaArgSerArgArgSerGlnGluThrGluValLeuTyrGlnLeuAlaHisThr 40
QY 139 CTGCCCTTTCGCGCGCGTGCAGCGCGCGACCTGCAGAGCGCTCCATCATCGCGCTCACA 198
DB 41 LeuProPheAlaArgGlyValSerAlaHisLeuAspLysAlaSerIleMetArgLeuThr 60
QY 199 ATCAGCTACCTGCATGTCACCGCCCTCTGCGCAGCAGGTGGAAGAGGGGAGGAGCCACT 258
DB 61 IleSerTyrLeuArgMetHisArgLeuCysAlaAlaGlyLysArgGlyArgAlaThr 80
QY 259 GGACGCTGTACTACTGAAGCCCTCGAGGGTTCGTCATGTGTACTCAGCGCGCGAGGAGA 318
DB 81 GlyArgLeuLeuProGluGlyProGlyGlyPheArgHisGlyThrHisArgArgGlyArg 100
QY 319 CATGGCTTACTCTCGGAATGTCAGCAGCAGTGGGCTCAGTCACTGGACCTCTGT 378
DB 101 HisGlyLeuProValGlyLysCysGlnGlnAlaProGlyProGlnSerValAspLeuCys 120
QY 379 TCCTCTCTCCCTGATACATAACCCCACTCCCTGGTACCAATTTCTCTGAGCTCATTTGA 438
DB 121 SerSerSerLeuIleHisAsnProThrProGlyThrAsnPheSerLeuGluLeuIleGly 140
QY 439 CACAGTATCTTTGATTTATCCATTCCTGTGACCAAGAGAACTTCAAGACGCCCTGCCCT 498
DB 141 HisSerIlePheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThr 160
QY 499 CCCAGCGCGAAGCTGTCAGAGAGAGCTGGAAGCCCCCAACAGAGCGCCACTTTTCCC 558
DB 161 ProArgProAsnLeuSerLysLysLysLeuGluAlaProThrGluArgHisPheSerLeu 180
QY 559 CGAATGAAGAGCAGCTCACAGCAGAGGCGCGACGCTCAACCTCAAAGCGCGCACCTGG 618
DB 181 ArgMetLysSerThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrp 200
QY 619 AAGTGCTGCATGCTCAGGACATATAGAGGCGCTACAAGCCCCCTGCACAGACTTCCCT 678
DB 201 LysValLeuHisCysSerGlyHisMetArgAlaTyrLysProAlaGlnThrSerPro 220
QY 679 GCGGAGCGCTCGTCCGAGCGCTCCCTCAATGCTGTGCTTATCTGTGAGCCATC 738
DB 221 AlaGlySerProArgSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIle 240
QY 739 CCCAGCTCCCTTCCACGATGTGTACTCTGTGGTCTTCCACAGGAGAGACTCCCATC 798
DB 241 ProGlnLeuProPheHisAspGlyAlaThrLeuGlyLeuProGlnGluLysThrProIle 260
QY 799 TCTACCTTATACCCCTCTTTTGAGAGGCACTACTTTGCTTGTCAAGAGGTGGCTGTT 858
DB 261 SerThrLeuPheThrProLeuTrpLysAlaLeuLeuCysLeuValLysArgTrpProVal 280
QY 859 CAGGTGCTCAGGGGAAAGGACTGAATCTCTCTCCCTCATGGTGTGTCTGGGCGCTT 918
DB 281 GlnValLeuGlnGlyLysGlyThrGluSerSerLeuProSerTrpValLeuTrpAlaLeu 300

QY 919 AACCGGAAAAATGTCTCTGGC 939
DB 301 AsnArgLysAsnCysProGly 307
RESULT 2
Q90X54
ID Q90X54 PRELIMINARY; PRT; 630 AA.
AC Q90X54;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypoxia-inducible factor 3 alpha (Fragment).
GN HIF3A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu Y.-Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.;
RT "Cloning and Characterization of a Third Hypoxia Inducible Factor,
HIF3-alpha.";
RL J. Biol. Chem. 0:0-0(1999).
DR EMBL; AF079153; AAF21782.1; JOINED.
DR EMBL; AF079140; AAF21782.1; JOINED.
DR EMBL; AF079141; AAF21782.1; JOINED.
DR EMBL; AF079142; AAF21782.1; JOINED.
DR EMBL; AF079143; AAF21782.1; JOINED.
DR EMBL; AF079144; AAF21782.1; JOINED.
DR EMBL; AF079145; AAF21782.1; JOINED.
DR EMBL; AF079146; AAF21782.1; JOINED.
DR EMBL; AF079147; AAF21782.1; JOINED.
DR EMBL; AF079148; AAF21782.1; JOINED.
DR EMBL; AF079149; AAF21782.1; JOINED.
DR EMBL; AF079150; AAF21782.1; JOINED.
DR EMBL; AF079151; AAF21782.1; JOINED.
DR EMBL; AF079152; AAF21782.1; JOINED.
DR MGD; MGI:1859778; Hif3a.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001067; Nuc_translocat.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTRNSLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
FT NON_TER 630
SQ SEQUENCE 630 AA; 69624 MW; 828EB2CB4E6D45B6 CRC64;

Alignment Scores:

Pred. No.: 7,86e-84 Length: 630
Score: 1045.00 Matches: 218
Percent Similarity: 85.88% Conservativity: 1
Best Local Similarity: 85.49% Mismatches: 5
Query Match: 51.99% Indels: 31
DB: 11 Gaps: 3

US-09-896-791B-2 (1-1100) x Q90X54 (1-630)

QY 43 AGTCTGAACACCGAGTGCAGGAGGAGAGTGCAGGAGCGCGCGCGCGCGCAGC 102
DB 7 ArgSerAsnThrGluLeuArgLysGluLysSerArgAspAlaAlaArgSerArgSer 26
QY 103 CAGGAGACGAGGTGTGTACACAGTGGCGCACACTCTGCCCCCTTGGCGGCGGTGACG 162
DB 27 GlnGluThrGluValLeuTyrGlnLeuAlaHisThrLeuProPheAlaArgGlyValSer 46
QY 163 GCGCACCTGCAGAACGCTCCATCATGCGCTCACATCAGCTACGTACGTGCGCACCGC 222
DB 47 AlaHisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrLeuArgMetHisArg 66

```
QY 223 CTCCTCGCAGCAGGTGGAA-----AAAGGGGAGAGCCACTGGAGCCGCTG 267
Db 67 LeuCysAlaAlaGly-GluTrpAsnGlnValGluLysGlyGluProLeuAspAlaCy 86
QY 268 CTACCTGAAGCCCTGGAGGTTTCGTATGCTACCTACCGCCGAGGAGACATGGCTTA 327
Db 86 sTyrLeuLysAlaLeuGluGlyPheValMetValLeuThrAlaGluGlyAspMetAlaTy 106
QY 328 CCTGTGCGAAATGTCAGCAGCAGCTGGGCTCAGTCAGTCAGTCAGTCAGTCAGTC 387
Db 106 rLeuSerGluAsnValSerLysHisLeuGlyLeuSerGln----- 119
QY 388 CTGATACATAACCCCACTCCTGGTACCAATTTCTCTCGAGCTCATTTGGACACAGTATC 447
Db 120 -----Leu--GluLeuIleGlyHisSerIle 127
QY 448 TTTGATTTATCCATCCCTGTGACCAAGAGGAGTTCAGAGCCCTGACCCCGAGCGCG 507
Db 128 PheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThrProArgPro 147
QY 508 AACCTGTCAAAGAAGAGCTGGAAGCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAG 567
Db 148 AsnLeuSerLysLysLeuGluAlaProThrGluArgHisPheSerLeuArgMetLys 167
QY 568 AGCAGCTCACAGCAGAGGGCCCTACAGGCCCTGACAGCCCTGACAGACTTCCCTGCCGGGAGC 687
Db 186 HisCysSerGlyHisMetArgAlaTyrLysProAlaGlnThrSerProAlaGlySer 207
QY 688 CCTCCTCGGAGCCCTCGCAATGCTGGTGTCTTATCTGTGAAGCCATCCCG----- 741
Db 208 ProArgSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIleProHisPro 227
QY 742 -----CAGCTCCCTTCCAGCATGGTGTCTACTCTG 771
Db 228 AlaSerLeuGluProProLeuGlyArgGlyAlaPheLeu 240
RESULT 3
Q92215 PRELIMINARY; PRT; 662 AA.
AC Q92215;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypoxia inducible factor three alpha.
GN HIF3A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN: [1]
RP SEQUENCE FROM N.A.
RA Gu Y.Z., Moran S.M., Hogenesch J.B., Wartman L., Bradfield C.A.;
RT "Molecular characterization and chromosomal localization of a third
RT alpha-class hypoxia induced factor subunit, HIF3alpha.";
RL Gene Expr. 0.0-0.1(1998).
DR EMBL; AF060194; AAC72734.1; -.
DR MGD; MGI:1859778; Hif3a.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001067; Nuc_translocat.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTRNSLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 662 AA; 73013 MW; 58740A1B6993D3B5 CRC64;
```

```
Alignment Scores:
Pred. No.: 7.93e-84 Length: 662
Score: 1045.00 Matches: 218
Percent Similarity: 85.88% Conservative: 1
Best Local Similarity: 85.49% Mismatches: 5
Query Match: 51.99% Indels: 31
DB: 11 Gaps: 3
US-09-896-791b-2 (1-1100) x Q92215 (1-662)
QY 43 AGTGCACACCGAGCTCGGGAAGGAGAGTCCGGGAGCGCGCGGAGCGGCGGAGC 102
Db 7 ArgSerAsnThrGluLeuArgLysGluLysSerArgAspAlaAlaArgSerArgSer 26
QY 103 CAGGAGAGCGAGGTCTGTACCAGCTGGCGCACACTGCTCCCTTTGCGCGCGCTGACG 162
Db 27 GlnGluThrGluValLeuIleGlnLeuAlaHisThrLeuProPheAlaArgGlyValSer 46
QY 163 GCGCACCTGGACAAGCCCTCCATCATGCGCTCACAATCAGCTACCTGCGCATGCCGCG 222
Db 47 AlaHisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrLeuArgMetHisArg 66
QY 223 CTCCTCGCAGCAGGTGGAA-----AAAGGGGAGAGCCACTGGAGCCCTG 267
Db 67 LeuCysAlaAlaGly-GluTrpAsnGlnValGluLysGlyGluProLeuAspAlaCy 86
QY 268 CTACCTGAAGCCCTGGAGGTTTCGTATGCTACTCTACCGCCGAGGAGACATGGCTTA 327
Db 86 sTyrLeuLysAlaLeuGluGlyPheValMetValLeuThrAlaGluGlyAspMetAlaTy 106
QY 328 CCTGTGCGAAATGTCAGCAGCAGCTGGGCTCAGTCAGTCAGTCAGTCAGTCAGTC 387
Db 106 rLeuSerGluAsnValSerLysHisLeuGlyLeuSerGln----- 119
QY 388 CTGATACATAACCCCACTCCTGGTACCAATTTCTCTGAGAGCTCATTTGGACACAGTATC 447
Db 120 -----Leu--GluLeuIleGlyHisSerIle 127
QY 448 TTTGATTTATCCATCCCTGTGACCAAGAGGAACTTCAAGAGCCCTGACCCCGAGCGCG 507
Db 128 PheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThrProArgPro 147
QY 508 AACCTGTCAAAGAAGAGCTGGAAGCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAG 567
Db 148 AsnLeuSerLysLysLeuGluAlaProThrGluArgHisPheSerLeuArgMetLys 167
QY 568 AGCAGCTCACAGCAGAGGGCCGACGCTCAACCTCAAGCGGCCACTTCCCTGCCGGGAGC 627
Db 168 SerThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrpLysValLeu 187
QY 628 CACTGCTCAGGACATATGAGGGCCCTACAGCCCTGACAGCCCTGACAGACTTCCCTGCCGGGAGC 687
Db 188 HisCysSerGlyHisMetArgAlaTyrLysProAlaGlnThrSerProAlaGlySer 207
QY 688 CCTCCTCGGAGCCCTCGCAATGCTGGTGTCTTATCTGTGAAGCCATCCCG----- 741
Db 208 ProArgSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIleProHisPro 227
QY 742 -----CAGCTCCCTTCCAGCATGGTGTCTACTCTG 771
Db 228 AlaSerLeuGluProProLeuGlyArgGlyAlaPheLeu 240
RESULT 4
OBWXAL PRELIMINARY; PRT; 632 AA.
ID OBWXAL
AC OBWXAL;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Inhibitory PAS domain protein.
GN IPAS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheng J.Q.;
RT "Cloning and characterization of human inhibitory PAS domain protein."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF463492; AAL69947.1; -;
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE; PS0112; PAS; 2.
SQ SEQUENCE 632 AA; 68964 MW; 9665B0AF3998F8EF CRC64;

Alignment Scores:
Pred. No.: 3.2e-81 Length: 632
Score: 1015.50 Matches: 212
Percent Similarity: 85.16% Conservative: 6
Best Local Similarity: 82.81% Mismatches: 8
Query Match: 50.52% Indels: 30
DB: 4 Gaps: 3

US-09-896-791B-2 (1-1100) x 08WXA1 (1-632)

QY 19 ATGGCGTTGGGCTGAGCGCGTGGAGTGCACACCGAGCTGGGAGGAGAGTCCGG 78
|||||
Db 1 MetAlaLeuGlyLeuGlnArgAlaArgSerThrThrGluLeuArgLysGluLysSerArg 20
QY 79 GAGCGCGCGCGAGCGCGAGCCAGGAGAGGAGTGTGTACCGCTGGCGCACACT 138
|||||
Db 21 AspaAlaArgSerArgSerGlnGluThrGluValLeuTyrGlnLeuAlaHisThr 40
QY 139 CTGCCCTTTTGGCGCGCTGACGCGCACCTGGACAGCGCTCCATCATGCGCTCACA 198
|||||
Db 41 LeuProPheAlaArgGlyValSerAlaHisLeuAspLysAlaSerIleMetArgLeuThr 60
QY 199 ATCAGCTACTCGGATGCACCGCTCTCGCAGCAGGT-----GGAAAA 243
|||||
Db 61 IleSerTyrLeuArgMetHisArgLeuCysAlaAlaGlyGluTrpAsnGlnValGly-Al 80
QY 244 AGGGGAGAGCCACTGAGCGCTGCTACCTGAGGCGCTGGAGGTTTGTCTATGCTACT 303
|||||
Db 80 aGlyGlyGluProLeuAspAlaCysTyrLeuLysAlaLeuGluGlyPheValMetValLe 100
QY 304 CACCGCGAGGAGACATGGCTTACCTGCGGAAAAATGTCAAGACACCTGGCGCTCAG 363
|||||
Db 100 uThrAlaGluGlyAspMetAlaTyrLeuSerGluAsnValSerLysHisLeuGlyLeuSe 120
QY 364 TCAGTGGAGCTGTGTTCTCTCCCTCCCTGATACATAACCCCACTCTGTTACCAATTTCTCT 423
|||||
Db 120 rGln-----Le 122
QY 424 CTGAGCTCATTTGGACACAGTATTTTATCCATCCCTGTGACCAAGAGGAACCTT 483
|||||
Db 122 u-GluLeuIleGlyHisSerIlePheAspPheIleHisProcysAspGlnGluGluLeu 141
QY 484 CRAAGCGCTTACCGCCCGGCGGACCTGTCAAAGAAAGTGAAGCCCAACAGAG 543
|||||
Db 142 GlnAspAlaLeuThrProGlnThrLeuSerArgArgLysValGluAlaProThrGlu 161
QY 544 CGGCACCTTTTCCGTGGATGAGAGCAGCTCACAGAGAGGGCGACGCTCAACCTC 603
|||||
Db 162 ArgCysPheSerLeuArgMetLysSerThrLeuThrSerArgGlyArgThrLeuAsnLeu 181
QY 604 AAAGCGGCCACCTGGAGGTGTGCTACTGCTCAGGACATATGAGGCGCTTACAGCCCTC 663
|||||
Db 182 LysAlaAlaThrTrpLysValLeuAsnCysSerGlyHisMetArgAlaTyrLysProPro 201
QY 664 GCACAGACTTCCCTCGCGGGAGCCCTCGCTCCGAGCCCTCCCTGCAATGCCTGTGCTT 723

Db 202 AlaGlnThrSerProAlaGlySerProAspSerGluProProLeuGlnCysLeuValLeu 221
|||||
QY 724 ATCTGTGAAGCCATCCCGCAGCTCCCTTCCACGATGGTGTCT 765
|||||
Db 222 IleCysGluAlaIlePro-----HisProGlySer 231
RESULT 5
Q96K34 PRELIMINARY; PRT; 632 AA.
AC Q96K34;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CDNA FLJ14819 fis, clone OVARC100241, moderately similar to hypoxia-inducible factor 1 alpha.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARIAN CARCINOMA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Kanehori K.;
RA "NEDO human cDNA sequencing project."
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK027725; BAB55324.1; -;
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00086; PAC; 1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 632 AA; 68934 MW; A19F1ED3D05E7A71 CRC64;

Alignment Scores:
Pred. No.: 7.23e-81 Length: 632
Score: 1011.50 Matches: 211
Percent Similarity: 84.77% Conservative: 6
Best Local Similarity: 82.42% Mismatches: 9
Query Match: 50.32% Indels: 30
DB: 4 Gaps: 3

US-09-896-791B-2 (1-1100) x Q96K34 (1-632)

QY 19 ATGGCGTTGGGCTGAGCGCGTGGAGTGCACACCGAGCTGGGAGGAGAGTCCGG 78
|||||
Db 1 MetAlaLeuGlyLeuGlnArgAlaArgSerThrThrGluLeuArgLysGluLysSerArg 20
QY 79 GAGCGCGCGCGAGCGCGCACCCAGGAGCGAGGTGTGTACCGCTGGCGCACACT 138
|||||
Db 21 AspaAlaArgSerArgSerGlnGluThrGluValLeuTyrGlnLeuAlaHisThr 40
QY 139 CTGCCCTTTTGGCGCGCTGACGCGCACCTGGACAGCGCTCCATCATGCGCTCACA 198
|||||
Db 41 LeuProPheAlaArgGlyValSerAlaHisLeuAspLysAlaSerIleMetArgLeuThr 60
QY 199 ATCAGCTACTCGGATGCACCGCTCTCGCAGCAGGT-----GGAAAA 243
|||||
Db 61 IleSerTyrLeuArgMetHisArgLeuCysAlaAlaGlyGluTrpAsnGlnValGly-Al 80
QY 244 AGGGGAGAGCCACTGGAGCGCTGCTACCTGAGGCGCTGGAGGTTTGTCTATGCTACT 303
|||||
Db 80 aGlyGlyGluProLeuAspAlaCysTyrLeuLysAlaLeuGluGlyPheValMetValLe 100
QY 304 CACCGCGAGGAGACATGGCTTACCTGCGGAAAAATGTCAAGACACCTGGCGCTCAG 363
|||||
Db 100 uThrAlaGluGlyAspMetAlaTyrLeuSerGluAsnValSerLysHisLeuGlyLeuSe 120

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=21134367; PubMed=11237857;
RA Kietzmann T., Cornesse Y., Brechtel K., Modaresi S., Jungermann K.;
RT "Perivenous expression of the mRNA of the three hypoxia-inducible
RT factor a-subunits HIF-1a, HIF2a and HIF3a in rat liver.";
RL Biochem. J. 354:531-537(2001).
DR EMBL: AJ277827; CAB96611.1;
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR001067; Nuc_translocat.
DR InterPro: IPR001610; PAC.
DR Pfam: PF00989; PAS; 2.
DR PRINTS: PR00785; NCTRNLOCATR.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 2.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 662 AA; 72887 MW; AC9672E340544010 CRC64;

Alignment Scores:
Pred. No.: 6.19e-80 Length: 662
Score: 1001.00 Matches: 211
Percent Similarity: 83.92% Conservative: 3
Best Local Similarity: 82.75% Mismatches: 10
Query Match: 49.80% Indels: 31
DB: 11 Gaps: 3

US-09-896-791B-2 (1-1100) x Q9Y2N7 (1-662)

QY 43 AGCTCGAACACGAGCTCGGAGGAGAGTCCGGGAGCGCGCCGCGAGCGCGCGCAGC 102
Db 7 ArgSerThrGluLeuArgLysGluLysSerArgAspAlaAlaArgSerArgSer 26

QY 103 CAGGACGAGGTGCTGACAGCTGGCGGCACACCTCTGCTTTCGCGCGCGCGTCAGC 162
Db 27 GlnGluThrGluValLeuTyrGlnLeuAlaHisThrLeuProPheAlaArgGlyValSer 46

QY 163 GCGCAGCTGGACAGCGCTCCATCATGCGCGCTCACATCAGTACCTGCGCATGACCGC 222
Db 47 AlaHisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrLeuArgMetHisArg 66

QY 223 CTCTGCGCAGCAGTGGAA-----AAAGGGGAGAGCCACTGGACCGCTG 267
Db 67 LeuLysAlaAlaGly-GluTrpAsnGlnValArgLysGluGlyGluProLeuAspAlaCy 86

QY 268 CTACCTGAAGCGCCTGGAGGGTTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
Db 86 stYrLeuLysAlaLeuGluGlyPheValMetValLeuThrAlaGluGlyAspMetAlaTy 106

QY 328 CTGTGCGGAATATGTCAGAACGACCTGGCGCTCATGCTAGTGAGCCTGTGCTCCTCC 387
Db 106 rLeuSerGluAsnValSerLysHisLeuGlyLeuSerGln----- 119

QY 388 CTGATACATAACCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
Db 120 -----Leu--GluLeuIleGlyHisSerIle 127

QY 448 TTTGATTTTATCCATCCCTGTGACCAAGAGACTTCAAGAGCGCTGACCCCGCGCGCG 507
Db 128 PheAspPheIleHisProCysAspGlnGluLeuGluAspAlaLeuThrProArgPro 147

QY 508 AACCTGTCAAAGAGAGTGGAGGCCCAACAGAGCGCCACTTTTCCTGCGAATGAG 567
Db 148 SerLeuSerLysLysSerGluAlaAlaThrGlyArgHisPheSerLeuArgMetLys 167

QY 568 AGCACGCTCAGCAGAGCGCGCAGCTCAACCTCAAGCGCCACCTGGAAGGTGCTG 627
Db 168 SerThrLeuThrSerArgGlyArgAlaLeuAsnLeuLysAlaAlaThrTrpLysValLeu 187

QY 628 CACTGCTCAGGACATATGAGGCGCTACAGCCCTTCACAGACTTCCCTCGCGGAGC 687
Db 188 HisCysSerGlyHisMetArgAlaTyrLysProProAlaGlnThrSerProAlaGlySer 207

QY 688 CCTCGCTCGAGCCTCCCTCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
Db 208 ProArgSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIleProHisPro 227

QY 742 -----CAGTCCCTCCCTCCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
Db 228 AlaSerLeuGluProProLeuGlnArgGlyAlaPheLeu 240

RESULT 8
Q9Y2N7 PRELIMINARY; PRT; 667 AA.

AC Q9Y2N7;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative homolog of hypoxia inducible factor three alpha
DE (Hypoxia-inducible factor-3 alpha).
GN HIF-3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Attix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
RT D19S412.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Hara S., Kobayashi C., Hamada J., Imura N.;
RT "Characterization of human hypoxia-inducible factor-3 alpha.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007193; RAD22668.1;
DR EMBL: AB054067; BAB69689.1;
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR001067; Nuc_translocat.
DR InterPro: IPR001610; PAC.
DR Pfam: PF00989; PAS; 2.
DR PRINTS: PR00785; NCTRNLOCATR.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 2.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 667 AA; 72404 MW; 67B8794FF9DCCF4B CRC64;

Alignment Scores:
Pred. No.: 2.68e-78 Length: 667
Score: 982.50 Matches: 205
Percent Similarity: 85.08% Conservative: 6
Best Local Similarity: 82.65% Mismatches: 7
Query Match: 48.88% Indels: 30
DB: 4 Gaps: 3

US-09-896-791B-2 (1-1100) x Q9Y2N7 (1-667)

QY 43 AGCTCGAACACGAGCTCGGAGGAGAGTCCGGGAGCGCGCCGCGAGCGCGCGCAGC 102
Db 7 ArgSerThrThrGluLeuArgLysGluLysSerArgAspAlaAlaArgSerArgSer 26

```

QY 103 CAGGACGGAGGTGCTGTACAGCTGGCGCACACTGCTCCCTTTCGCGCGCGCTCAGC 162
Db 27 GlnGluThrGluValLeuTyrGlnLeuAlaHisThrLeuProPheAlaArgGlyValSer 46
QY 163 GCGACCTGGACAAGCCTCCATCATGCGCTCAACATCAGTACCTGGCATGCGCACCGC 222
Db 47 AlaHisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrLeuArgMetHisArg 66
QY 223 CTCTCGGACGAGGTG-----GGAAAAAGGGGAGAGCCACTGGACGCGCTG 267
Db 67 LeuGlyAlaAlaGlyGluTrpAsnGlnValGly--AlaGlyGlyGluProLeuAspAlaC 86
QY 268 CTACCTGAAGCCCTCGAGGGTTTCGTATGCTACTACCGCGGAGGAGACATGGCTTA 327
Db 86 styrLeuLysAlaLeuGluGlyPheValMetValLeuThrAlaGluGlyAspMetAlaTy 106
QY 328 CCTGTGGAATGTCAGAACCACTGGCCCTCAGTCAGTGGAGCTCTGTTCCTCCCTCC 387
Db 106 rLeuSerGluAsnValSerLysHisLeuGlyLeuSerGln----- 119
QY 388 CTGATACATAACCCCACTCCTGGTACCAATTTCTCTCTGGAGCTCATTTGGACACAGTATC 447
Db 120 -----Leu--GluLeuIleGlyHisSerIle 127
QY 448 TTTGATTTTATCCATCCTGTGACCAAGAGGAACCTTCAAGAGCGCCTGACCCCGAGCGG 507
Db 128 PheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThrProGlnGln 147
QY 508 AACCTGTCNAGAAGAAGCTGGAAGCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAG 567
Db 148 ThrLeuSerArgLysValGluAlaProThrGluArgCysPheSerLeuArgMetLys 167
QY 568 AGCAGCTCACAGAGGCGCGAGCTCAACCTCAAGCGCGCCACCTGGAGGTGCTG 627
Db 168 SerThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrpLysValLeu 187
QY 628 CACTGCTCAGGACATATGAGGCGCTACAGCGCCCTCGACAGACTTCCCTCGCGGAGC 687
Db 188 AsnCysSerGlyHisMetArgAlaTyrLysProProAlaGlnThrSerProAlaGlySer 207
QY 688 CCTCGCTCGGACCTCCCTGCAATGCGCTGTGTGCTTATCTGTGAAGCCATCCCGCAGTCT 747
Db 208 ProAspSerGluProLeuGlnCysLeuValLeuIleCysGluAlaIlePro----- 225
QY 748 CCTTCCAGAGTGTGCT 765
Db 226 -----HisProGlySer 229

RESULT 9
Q9HAI2
ID Q9HAI2 PRELIMINARY; PRT; 648 AA.
AC Q9HAI2:
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical 70.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
RA Niimiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK021653; BAB13865.1;

```

```

DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
SQ SEQUENCE 648 AA; 69994 MW; EBFEC744BC3F148E CRC64;

Alignment Scores:
Pred. No.: 4,82e-58 Length: 648
Score: 753.50 Matches: 158
Percent Similarity: 82.00% Conservative: 6
Best Local Similarity: 79.00% Mismatches: 6
Query Match: 37.49% Indels: 30
DB: 4 Gaps: 3

US-09-896-791b-2 (1-1100) x Q9HAI2 (1-648)
QY 187 ATGCGCCCTCACATCAGCTACCTGGCATGACCGCCCTTGGCGAGCAGGT----- 237
Db 1 MetArgLeuThrIleSerTyrLeuArgMetHisArgLeuCysAlaAlaGlyGluTrpAsn 20
QY 238 -----GGAAAAAGGGGAGAGCCACTGGACGCGCTGCTACCTGAAGGCGCTGGAGGTTT 291
Db 21 GlnValGly--AlaGlyGlyGluProLeuAspAlaCysTyrLeuLysAlaLeuGluGlyPh 40
QY 292 CGTCATGTTACTCACCGCGAGGAGACATGGCTTACCTGTCGGAATAATGTCAGCAAGCA 351
Db 40 eValMetValLeuThrAlaGluGlyAspMetAlaTyrLeuSerGluAsnValSerLysH 60
QY 352 CTTGGGCTCAGTCAGTGGAGCTCTGTTCTCTCTCCCTGATACATAACCCCACTCTCTGTT 411
Db 60 sLeuGlyLeuSerGln----- 65
QY 412 ACCAATTTCTCTCGAGCTCATTTGGACACATATCTTTGATTTTATCCATCCCTGTGAC 471
Db 66 -----Leu--GluLeuIleGlyHisSerIlePheAspPheIleHisProCysAsp 81
QY 472 CAAGAGGACTTCAAGAGCCCTGACCCCGAGCGCAACCTGTCGAAGAAGAGCTGGAA 531
Db 82 GlnGluLeuGlnAspAlaLeuThrProGlnGlnThrLeuSerArgArgLysValGlu 101
QY 532 GCCCAACAGAGCGGCACCTTTTCCCTGCGAATGAAGAGCACGCTCACACAGAGGCGC 591
Db 102 AlaProThrGluArgCysPheSerLeuArgMetLysSerThrLeuThrSerArgGlyArg 121
QY 592 AGCTCAACCTCAAGCGCCACCTGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
Db 122 ThrLeuAsnLeuLysAlaAlaThrTrpLysValLeuAsnCysSerGlyHisMetArgAla 141
QY 652 TACAAGCCCTCGACAGACTTCCCTGCGGAGCGCCCTGCTCGAGCTCCCTGCA 711
Db 142 TyrLysProProAlaGlnThrSerProAlaGlySerProAspSerGluProLeuGln 161
QY 712 TGCCTGGTGTATCTGTGAAGCCATCCCGCAGCTCCCTTCCACGATGGTGT 765
Db 162 CysLeuValLeuIleCysGluAlaIlePro-----HisProGlySer 175

RESULT 10
Q9HAM5
ID Q9HAM5 PRELIMINARY; PRT; 199 AA.
AC Q9HAM5:
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Hypothetical 21.6 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

```



```
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK021421; BAB13819.1; -.
SQ SEQUENCE 199 AA; 21597 MW; 646CB066A09B47C3 CRC64;

Alignment Scores:
Pred. No.: 6,250-51 Length: 199
Score: 672.00 Matches: 141
Percent Similarity: 64.19% Conservative: 6
Best Local Similarity: 61.57% Mismatches: 14
Query Match: 33.43% Indels: 68
DB: 4 Gaps: 2

US-09-896-791B-2 (1-1100) x Q9H4M5 (1-199)
QY 83 CGGCCGCGAGCGCGGCGAGGAGCGAGGAGTGTGTCACGCTGGGCGACACACTCTGC 142
Db 2 ArgProAlaAlaGlyAlaAlaArgArgProArgCysCysThrSerTrpLeuThrArgCys 21
QY 143 CTTTGGCGCGCGGCTGACGCGGCGACCTGGACAAAGCGCTCCATCATGCGGCTCACAAATCA 202
Db 22 ProSerProAlaAlaSerAlaProThrTrpThrArgProLeuSerCysAlaSerProSer 41
QY 203 GCTACCTGCGCATGTCACCGCTCTGCGCAGCGAGTGGAAAAAGGAGGAGCGACCTGCGAC 262
Db 42 AlaThrCysAlaCysThrAlaSerAlaProGln----- 52
QY 263 GCCTGTACTGTAAGCGCCTGGAGGCTTTCGTATCATGTACTACCGCGGAGGAGACATG 322
Db 52 ----- 52
QY 323 GCTTACTCTGCGAAATATGTCACAGACACCTGGGCGCTCAGTCAGTGGACCTGTGTCCT 382
Db 52 ----- 52
QY 383 CTTCCCTGTACATAACCCGACTCTCTGGTACCAATTTCTCTGGAGCTCATTTGGACACA 442
Db 53 -----Leu-GluLeuIleGlyHis 59
QY 443 GTATCTTTGATTTATTCATTCCTGTGACAGAGAACTTCAAGACGCGCCTGACCCCA 502
Db 59 erillepheaspheilleHisProCysAspGlnGluLeuGlnAspAlaLeuThrProG 79
QY 503 GCGCGAAGCTGTCAAGAAGAAGCTGGAAGCCCAACAGAGCGCCACTTTTCCCTCGGA 562
Db 79 InGlnThrLeuSerArgArgLysValGluAlaProThrGluArgCysPheSerLeuArgM 99
QY 563 TGAAGAGCAGCTCACAGCAGAGCGCGACGCTCAACCTCAAGCGCGCACCTGGAAGG 622
Db 99 etlysserThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrpLysV 119
QY 623 TGTGTCAGCTGCTCAGACATATGAGGCGCTACAGCCCGCTGCACAGACTTCCCTCGCG 682
Db 119 alLeuAsnCysSerGlyHisMetArgAlaTyrLysProProAlaGlnThrSerProAlaG 139
QY 683 GGAGCCCTGCTCGGAGCTCCCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
Db 139 LysSerProAspSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIlePro- 158
QY 743 AGCTCCCTTCCACGATGTGCT 765
Db 159 -----HisProGlySer 162

RESULT 11
Q9CYA8 PRELIMINARY; PRT; 258 AA.
AC Q9CYA8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
```

```
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypoxia inducible factor 1, alpha subunit.
GN HIF1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Caasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Ouackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK017853; BAB30975.1; -.
DR MGD; MGI:106918; Hif1a.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001067; Nuc_translocat.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00989; PAS; 1.
DR PRINTS; PR00785; NCTRNSLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00091; PAS; 1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 258 AA; 29703 MW; C42D1CD68117C552 CRC64;
```

Alignment Scores:

Pred. No.:	9,82e-46	Length:	258
Score:	613.50	Matches:	129
Percent Similarity:	67.49%	Conservative:	35
Best Local Similarity:	53.09%	Mismatches:	51
Query Match:	30.52%	Indels:	28
DB:	11	Gaps:	3

US-09-896-791B-2 (1-1100) x Q9CYA8 (1-258)

```
QY 34 CAGCGCGTGTGAGTGCAGACACCGAGTGGGAGAGAGAGTGGGAGCGGCGCCGAGC 93
Db 9 GluLysLysLysMetSerSerGluArgArgLysGluLysSerArgAspAlaLaArgSer 28
QY 94 CGCGCGCAGCAGGAGACGAGGTGTGTACAGCTGGCGCACACTCTGCTTGGCGCC 153
Db 29 ArgArgSerLysGluSerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHis 48
QY 154 GCGCTCAGCGCGCACCTGGCAAGGCTCCATCATCGCTGCATCATCATCATCATCAT 213
Db 49 AsnValSerSerHisLeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArg 68
QY 214 ATGCACCGCTCTGCGCAGCAGCTGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 258
Db 69 ValArgLysLeuLeuAspAlaGlyGly-LeuAspSerGluAspGluMetLysAlaGlnMe 88
QY 259 GGACGCGCTGCTACCTCAAGCGCCTGGAGGTTTCGTTCATGCTGCTGCTGCTGCTG 318
Db 88 taspCysPheTyrLeuLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAs 108
```



```
QY 319 CATGCTTACCTGTCGGAATAATGTGACGAGCACTGGCGCTCAGTCAGTGACCTCTGT 378
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 pMetValTyrIleSerAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124

QY 379 TCCTCTCCCTGATACATACCCCACTCCTGGTACCAATTTCTCTCTGGAGCTATTGGA 438
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 -----Phe--GluLeuThrGly 129

QY 439 CACAGTATCTTTGATTTCATCCCTGTGACCAAGAGAACTTCAAGACGCCCTGACC 498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 HisSerValPheAspPheThrHisProCysAspHisGluMetArgGluMetLeuThr 149

QY 499 CCCAGGCCCACTGCTCAAGAAGAGCTGGAAGCCCAACAGAGCGCCACTTTTCCCTG 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 150 HisArgAsnGlyProValArgLysGlyLysGluLeuAsnThrGlnArgSerPheLeu 169

QY 559 CGAATGAGAGACGACCTCACAGAGAGGGCGCAGCTCAACCTCAAGCGGCCACCTGG 618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 170 ArgMetLysCysThrLeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrp 189

QY 619 AAGGTGCTGCACTGCTCAGGACATATAGGGCTACAAGCCCTGCACAGACTTCCCT 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 LysValLeuHisCysThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGln 209

QY 679 GCCGGGAGCCCTCGCTCCGAGCCTCCCTGCAATGCCTGGTGTATCTGTGAAGCCATC 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 210 CysGly-----TyrLysLysProPrometThrCysLeuValLeuIleCysGluProIle 227

QY 739 CCC 741
    |||||
Db 228 Pro 228

RESULT 12
Q8R385
ID Q8R385 PRELIMINARY; PRT; 836 AA.
AC Q8R385;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypoxia inducible factor 1, alpha subunit.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026139; AAH26139.1; -.
SQ SEQUENCE 836 AA; 93469 MW; 8A7B21B0F13E84D5 CRC64;
```

```
Alignment Scores:
Pred. No.: 1.23e-45 Length: 836
Score: 613.50 Matches: 129
Percent Similarity: 67.49% Conservative: 35
Best Local Similarity: 53.09% Mismatches: 51
Query Match: 30.52% Indels: 28
DB: 11 Gaps: 3

US-09-896-791B-2 (1-1100) x Q8R385 (1-836)

QY 34 CAGCGCGTGAGTCCGACACCCAGCTGCGGAAGGAGAGTCCGCGGAGCGCGCCGAGC 93
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9 GluLysLysLysMetSerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSer 28

QY 94 CGCGCAGCCAGGAGAGGTGCTGTACCAAGTGGCGCACACCTCTGCCCTTTGCGGCG 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29 ArgArgSerLysGluSerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHis 48

QY 154 GCGTTCAGCGCCAGCTGCAGAGCCCTCATCATCGCGCTCACAAATCAGCTACCTGGCG 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 49 AsnValSerSerHisLeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArg 68

QY 214 ATGCACCGCCTCTGCGCAGCAGTGGAAAAAGGGGAGAG-----CCACT 258
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 69 ValArgLysLeuLeuAspAlaGlyGly-LeuAspSerGluAspGluMetLysAlaGlnMe 88

QY 259 GGACGCCCTGCTACCTGAAGGCCCTGGAGGTTTGTGTCATGGTACTACCGCGGAGGAGA 318
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88 tAspCysPheTyrLeuLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAs 108

QY 319 CATGCTTACCTGTCGGAATAATGTGACGAGCACTGGCGCTCAGTCAGTGCACCTCTGT 378
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 pMetValTyrIleSerAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124

QY 379 TCCTCTCCCTGATACATACCCCACTCCTGGTACCAATTTCTCTCTGGAGCTCATTTGA 438
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 -----Phe--GluLeuThrGly 129

QY 439 CACAGTATCTTTGATTTCATCCCTGTGACCAAGAGAACTTCAAGACGCCCTGACC 498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 HisSerValPheAspPheThrHisProCysAspHisGluMetArgGluMetLeuThr 149

QY 499 CCCAGGCCCACTGCTCAAGAAGAGCTGGAAGCCCAACAGAGCGCCACTTTTCCCTG 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 150 HisArgAsnGlyProValArgLysGlyLysGluLeuAsnThrGlnArgSerPheLeu 169

QY 559 CGAATGAGAGACGACCTCACAGAGAGGGCGCAGCTCAACCTCAAGCGGCCACCTGG 618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 170 ArgMetLysCysThrLeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrp 189

QY 619 AAGGTGCTGCACTGCTCAGGACATATAGGGCTACAAGCCCTGCACAGACTTCCCT 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 LysValLeuHisCysThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGln 209

QY 679 GCCGGGAGCCCTCGCTCCGAGCCTCCCTGCAATGCCTGGTGTATCTGTGAAGCCATC 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 210 CysGly-----TyrLysLysProPrometThrCysLeuValLeuIleCysGluProIle 227

QY 739 CCC 741
    |||||
Db 228 Pro 228

RESULT 13
Q9XTA5
ID Q9XTA5 PRELIMINARY; PRT; 823 AA.
AC Q9XTA5;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Hypoxia-inducible factor-1 alpha.
GN HIF-1 ALPHA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hara S.; Kobayashi C.; Imura N.;
RL "Molecular cloning of cDNAs encoding hypoxia-inducible factor (HIF)-
RL 1alpha and -2alpha of bovine arterial endothelial cells.";
RL Biochim. Biophys. Acta 1445:237-243(1999).
DR EMBL; AB018398; BAA78675.1; -.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001321; HypoxindfIA.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR01080; HYPOXIAFIA.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
```

DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 823 AA; 92127 MW; 12674E467A61B1A1 CRC64;

Alignment Scores:

Pred. No.: 1-84e-45 Length: 823
Score: 611.50 Matches: 130
Percent Similarity: 68.49% Conservative: 33
Best Local Similarity: 54.62% Mismatches: 47
Query Match: 30.42% Indels: 28
DB: 3 Gaps: 3

US-09-896-791B-2 (1-1100) x Q9YIB9 (1-823)

QY 49 AACACCGAGCTCGAGGAGAGAGTCCGGGACGCGCGCGCGCGCGCGCGCGCGCGAGG 108
Db 14 SerSerGluArgArgLysGluLysSerArgAlaAlaArgSerArgArgSerLysGlu 33
QY 109 ACGAGGTGCTGACAGCTGGCGCACACTCTGCCCTTTCCGCGCGCGGTGACGGCGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 53
QY 169 CTGCACAGCGCTCCATCATCGCTCCACATCAGCTACCTGCGCATGCGACCGCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
QY 229 GCACGAGGTGGAAAGGGGGAGAG-----CCACTGGACGCGCTGCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspGluMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCGCTGGAGGGTTTCATGTTACTACCGCGGAGGAGACATGCTTACTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GAAATATGTCAGCAAGCACTGGCGCTCAGTCAGTGGACCTGTTCTCTCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATACCCCACTCCTGGTACCAATTTCTCTGGAGCTCATTTGGACACAGATATCTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCCTGTGACCAAGAGGAACTTCAAGACGCGCTGACCGCGGCGGAGCACTG 513
Db 135 PheThrHisProCysAspHisGluMetArgLysGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAAGCTGGAAGCCCAACAGACGCGCACTTTTCCCTGCGGAATGAAGACGACG 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACGACGAGGGCGGCGCTCAACCTCAAGCGCGCCACCTGGAAGGTGCTGCACTGC 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATAGGCGCTACAGCGCCCTGCGACAGACTTCCCTCGCGGGAGCCCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnSerGlnCysGly-----Tyr 212
QY 694 TCGAGCTCCCTCGCAATGCTGGTGTATCTGTTGAGCCATCCCC 741
Db 213 LysLysProPromethrCysLeuValLeuIleCysGluProIlePro 228

RESULT 14

Q9YIB9
ID Q9YIB9 PRELIMINARY; PRT; 811 AA.
AC Q9YIB9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypoxia-inducible factor-1 alpha.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA Takahashi T.;

RT "Molecular cloning and expression of an avian cDNA for hypoxia-inducible factor-1 alpha in embryonic ventricular myocytes.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB013746; BAA34234.2; -;
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.

DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.

DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 811 AA; 90542 MW; D14CD9FC9BF064CB CRC64;

Alignment Scores:

Pred. No.: 4-14e-45 Length: 811
Score: 607.50 Matches: 136
Percent Similarity: 67.36% Conservative: 25
Best Local Similarity: 56.90% Mismatches: 51
Query Match: 30.22% Indels: 28
DB: 13 Gaps: 3

US-09-896-791B-2 (1-1100) x Q9YIB9 (1-811)

QY 43 AGTTCGAACACCGAGCTCGGGAAGAGAGTCCGGGACGCGCGCGCGCGCGCGAGC 102
Db 12 ArgIleSerSerGluArgArgLysGluLysSerArgAlaAlaArgCysArgSer 31
QY 103 CAGGACGACGAGTGTGTACCACTGGCGCACACTCTGCCCTTGGCGCGCGCTCAGC 162
Db 32 LysGluSerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisThrValSer 51
QY 163 GCCACCTGGACAAGCGCTCATCATCGCGCTCACATCAGCTACCTGCGCATGCGCGC 222
Db 52 AlaHisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrLeuArgMetArgLys 71
QY 223 CTCTGCGCAGAGTGGAAAAGGGGGAGAG-----CCACTGGACGCGCTG 267
Db 72 LeuLeuAspAlaGly-GluLeuGluThrGluAlaAsnMetGluLysGluLeuAsnCysPh 91
QY 268 CTACCTGAAGGCCCTGGAGGTTTTCATGTTACTACCGCGCGGAGGAGACATGGCTTA 327
Db 91 eTyrLeuLysAlaLeuAspGlyPheValMetValLeuSerGluAspGlyAspMetIleTyr 111
QY 328 CTGTGCGAAAATGTGAGCAAGCACCTGGCGCTCAGTCAGTG--GACCTCTGTTCCTCCTC 386
Db 111 rMetSerGluAsnValAsnLysCysMetGlyLeuThrGlnPheAsp----- 126
QY 387 CCTGATACATACCCCACTCCTGGTACCAATTTCTCTGAGCTCATTTGGACACAGATAT 446
Db 127 -----LeuThrGlyHisSerVa 132
QY 447 CTTTGATTTTATCCATCCCTGTGTCACCAAGAGGAACCTTCAAGACGCGCTGACCCCGAGGCC 506
Db 132 lPheAspPheThrHisProCysAspHisGluLeuArgLysGluMetLeuThrHisArgAs 152
QY 507 GAACCTGTCAAGAAGAAGCTGGAAGCCCAACAGACGCGCACTTTTCCCTGCGAATGAA 566
Db 152 nGlyProValLysLysGlyLysGluGlnAsnThrGluArgSerPhePheLeuArgMetLys 172
QY 567 GAGCAGCTTCACACGAGGCGCGCGCTCAACCTCAAGCGCGCCACCTTGAAGTGTCT 626
Db 172 sCysThrLeuThrSerArgGlyArgThrValAsnIleLysSerAlaThrTrpLysValLe 192
QY 627 GCACCTCTCAGGACATATGAGGCGCTTACAGCGCCCTGACAGACTTCCCTCGCGGGAG 686
Db 192 uHisCysThrGlyHisIleArgValTyrAspThrCysAsnAsnGlnThrHisCysGly-- 211

QY 687 CCCTGGCTCCGAGCCTCCCTGCAATGCTGCTGCTATCTGTGAGCCATCCCC 741
 Db 212 ---TyrLysLysProProMetThrCysLeuValLeuIleCysGluProIlePro 228
 RESULT 15
 Q96PT9 ID Q96PT9 PRELIMINARY; PRT: 735 AA.
 AC Q96PT9; 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypoxia-inducible factor 1 alpha variant.
 GN HIF1A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Tanaka S., Sugimachi K.;
 RT "Hypoxia-Inducible factor-1 alpha variant isolated from human liver tissue.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB073325; BAB70608.1; -;
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS_domain.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 2.
 DR SMART; SM00086; PAC; 1.
 DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
 SQ SEQUENCE 735 AA; 82746 MW; 34DD604FB4E4418E CRC64;

Alignment Scores:

Pred. No.: 4,98e-45 Length: 735
 Score: 606.50 Matches: 129
 Percent Similarity: 68.07% Conservative: 33
 Best Local Similarity: 54.20% Mismatches: 48
 Query Match: 30.17% Indels: 28
 DB: 4 Gaps: 3

US-09-896-791B-2 (1-1100) x Q96PT9 (1-735)

QY 49 ACACCGAGCTCGGAGGAGAGTCCGCGGAGCGCGCGCGCGCGCGAGCAGGAG 108
 Db 14 SerSerGluArgArgLysLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
 QY 109 ACGGAGGTCTCTACAGCTGGCGCACACTCTGCCCTTTTGGCGCGCGCTCAGCGGCAC 168
 Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerHis 53
 QY 169 CTGGACAGGCTCCATCATGCGCTCACAATCAGCTACCTGCGATGCGCGCTCTGC 228
 Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeu 73
 QY 229 GCAGCAGTGGGAAAGGGGGAGAG-----CCACTGGAGCGCTGTACCT 273
 Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnPheTyrLe 93
 QY 274 GAAGCCCTGGAGGTTTCGTCTGCTATGCTACCTACCGCGGAGGAGACATGGCTTACCTGTC 333
 Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
 QY 334 GGAATATGCAGCAGCAGCTGGGCTCAGTCAGTGGACCTCTGTCTCTCCCTCCCTGATA 393
 Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
 QY 394 CATAACCCCATCTCTGGTACCAATTCTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
 Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134

Search completed: December 2, 2002, 20:39:17
 Job time : 70 secs

QY 454 TTTATCCATCCTGTGACCAAGAGGAACTTCAAGACGCGCCTGACCCCGAGCGCAACCTG 513
 Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
 QY 514 TCAAAGAAGAAGCTGGAAGCCCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGACACG 573
 Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
 QY 574 CTCACGACGAGCGGCGCACGCTCAACCTCAAAGCGCGCCACTGGAGAGTGTGCACTGC 633
 Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrIrrLysValLeuHisCys 194
 QY 634 TCAGGACATATGAGGCGCTTACAAGCCCCCTGCACAGACTTCCCTCGCGGAGCCCTCGC 693
 Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
 QY 694 TCCGAGCCTCCCTGCAATGCGCTGGTGTATCTGTGAAGCCATCCCC 741
 Db 213 LysLysProProMetThrCysLeuValLeuIleCysGluProIlePro 228

THIS PAGE BLANK (USP10)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 2, 2002, 18:23:15 ; Search time 55.5 Seconds
(without alignments)
5282.007 Million cell updates/sec

Title: US-09-896-791B-2
Perfect score: 2010
Sequence: 1 gaattcgacagggccat.....aaaaaaaaaacatcgccgc 1100

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=rlp
-MODE=frame+n2p.model -DEV=rlp
-Q/cgn2_1/USPTO.spool/US09896791/runat_02122002_152657_5564/app_query.fasta_1.1287
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896791 @CGL_1.1.42 @runat_02122002_152657_5564 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEVTIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1636	81.4	307	23	AAU75902	Mouse inhibitory P
2	1045	52.0	662	20	AAV06295	Mouse transcriptio
3	1011.5	50.3	632	22	AA93326	Human protein sequ
4	982.5	48.9	667	23	AAE24222	Human HIF-3 protei
5	978	48.7	407	23	ABG66737	Human novel polype
6	753.5	37.5	648	22	AA94934	Human protein sequ
7	672	33.4	199	22	AA93710	Human protein sequ
8	610.5	30.4	810	23	ABB57270	Mouse ischaemic co
9	606.5	30.2	245	21	AAV94628	HIF-lalpha variant
10	606.5	30.2	330	21	AAV94628	HIF-lalpha variant
11	606.5	30.2	642	21	AAV84168	A variant of human
12	606.5	30.2	652	21	AAV94629	HIF-lalpha variant
13	606.5	30.2	669	21	AAV84167	A variant of human
14	606.5	30.2	697	21	AAV84166	A variant of human
15	606.5	30.2	701	21	AAV84173	A variant of human
16	606.5	30.2	710	21	AAV84172	A variant of human
17	606.5	30.2	724	21	AAV84171	A variant of human
18	606.5	30.2	749	21	AAV84170	A variant of human
19	606.5	30.2	789	21	AAV84169	A variant of human
20	606.5	30.2	813	21	AAV94636	HIF-lalpha variant
21	606.5	30.2	826	18	AAW06557	Human hypoxia indu
22	606.5	30.2	826	19	AAW80418	Amino acid sequenc
23	606.5	30.2	826	20	AAV06289	Human transcriptio
24	606.5	30.2	826	21	AAV94640	Human hypoxia-indu
25	606.5	30.2	826	21	AAV69407	A wild type human
26	606.5	30.2	826	22	AAV76854	Human lung tumour
27	606.5	30.2	826	23	AAE24212	Human HIF-lalpha p
28	606.5	30.2	826	23	AAV77602	Human hypoxia-indu
29	606.5	30.2	826	23	AAV77606	Human hypoxia-indu
30	606.5	30.2	826	23	AAV77607	Human hypoxia-indu
31	606.5	30.2	826	23	AAV77608	Human hypoxia-indu
32	606.5	30.2	826	23	AAV77609	Human hypoxia-indu
33	606.5	30.2	826	23	AAV77610	Human hypoxia-indu
34	606.5	30.2	826	23	AAV77611	Human hypoxia-indu
35	606.5	30.2	826	23	AAV77612	Human hypoxia-indu
36	606.5	30.2	826	23	AAV77613	Human hypoxia-indu
37	606.5	30.2	826	23	AAV77615	Human hypoxia-indu
38	606.5	30.2	826	23	AAV77616	Human hypoxia-indu
39	606.5	30.2	826	23	AAV77617	Human hypoxia-indu
40	606.5	30.2	826	23	AAV77618	Human hypoxia-indu
41	606.5	30.2	826	23	AAV77619	Human hypoxia-indu
42	606.5	30.2	826	23	AAV77620	Human hypoxia-indu
43	606.5	30.2	826	23	AAU85509	Clone #19080 of lu
44	606.5	30.2	826	23	AAU69409	Lung small cell ca
45	583	29.0	875	19	AAW37098	Murine endothelial

ALIGNMENTS

RESULT 1
AAU75902
ID AAU75902 standard; Protein; 307 AA.
XX
AC AAU75902;
XX
DT 08-MAY-2002 (first entry)
XX
DE Mouse inhibitory PAS domain (IPAS) protein.
XX
KW Inhibitory PAS domain protein; IPAS; cardiant; vasotropic;
KW cardiovascular; cerebroprotective; ophthalmological; HIF-lalpha;
KW hypoxia-induced factor lalpha; vascular endothelial growth factor;
KW angiogenesis; ischaemic cardiovascular lesion; stroke;
KW diabetic microvascular disease; tumour; mouse.
XX
OS Mus musculus.
XX
PN WO200202609-A1.

10-JAN-2002.
 19-JUN-2001; 2001WO-SE01387.
 06-JUL-2000; 2000SE-0002551.
 (BIOV-) BIOVITRUM AB.
 Berkenstam A, Bertilsson G, Poellinger L;
 WPI; 2002-164523/21.
 N-PSDB; ABK14502.
 New nucleic acid encoding inhibitory PAS domain protein, useful for
 identifying specific inhibitors for treating e.g. angiogenesis or
 tumour growth
 Claim 3; Fig 1; 44pp; English.
 The invention describes an isolated nucleic acid encoding the
 biologically active inhibitory PAS domain protein or its functionally
 equivalent modifications. IPAS forms a non-functional heterodimeric
 complex with HIF-1alpha (hypoxia-induced factor 1alpha), impairing
 interaction between HIF-1alpha and hypoxia-response elements in genes,
 e.g. the gene for vascular endothelial growth factor, so contributes to
 control of hypoxic signalling. The nucleic acid and its encoded
 polypeptides, are used to identify agents that activate expression of
 the gene or stimulate activity of the protein. These agents are useful
 for inhibiting angiogenesis, particularly where associated with ischaemic
 cardiovascular lesions, stroke or diabetic microvascular diseases, and
 tumour growth. This is the amino acid sequence of the mouse inhibitory
 PAS domain protein (IPAS), described in the method of the invention.
 Sequence 307 AA;

Alignment Scores:

Pred. No.: 9,75e-139 Length: 307
 Score: 1636.00 Matches: 307
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 81.39% Indels: 0
 DB: 23 Gaps: 0

US-09-896-791B-2 (1-1100) x AAU75902 (1-307)

QY 19 ATGGCGTTGGCGTGCAGCGCTGAGTGCAGACACCGAGTGCAGAGGAGAGTCCGG 78
 Db 1 MetAlaLeuGlyLeuGlnArgValArgSerAsnThrGluLeuArgLysGluLysSerArg 20
 QY 79 GACGCGCCGCGCAGCGCGCAGCAGGAGCGAGTGTGTACACGCTGGCGCACACT 138
 Db 21 AsplaalaarGserArgSerGlnGluThrGluValLeuYrGlnLeuAlaHisThr 40
 QY 139 CTGCCCTTTGCGCGCGCGTGCAGCGCACCTGGACAGGCGCTCCATCGCGCTCACA 198
 Db 41 LeuProPheAlaArgGlyValSerAlaHisLeuAspLysAlaSerIleMetArgLeuThr 60
 QY 199 ATCAGCTACCTGCAGTGCACCGCTCTGCGCAGCAGTGTGAAAGGGGAGAGCCACT 258
 Db 61 IleSerTyrLeuArgMetHisArgLeuLysAlaAlaGlyGlyLysArgGlyArgAlaThr 80
 QY 259 GGACCGCTGTACTGAAGGCGCTGGAGGTTTCGTATGCTACTACCGCCGAGGAGAGA 318
 Db 81 GlyArgLeuLeuProGluGlyProGlyGlyPheArgHisGlyThrHisArgArgGlyArg 100
 QY 319 CATGGCTTACCTGTCCGAAAATGTACAGACACCTGGCGCTCAGTCAGTGGACCTCTGT 378
 Db 101 HisGlyLeuProValGlyLysCysGlnGlnAlaProGlyProGlnSerValAspLeuLys 120
 QY 379 TCCTCTCTCTGATACATACCCCACTCTGGTACCAATTTCTCTCTGGAGCTCATTTGGA 438
 Db 121 SerSerSerLeuIleHisAsnProThrProGlyThrAsnPheSerLeuGluLeuIleGly 140

QY 439 CACAGTATCTTTGATTATCCATCCCTGTGTGACCAAGAGGAACTTCAAGAGCCCTGACC 498
 Db 141 HisSerIlePheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThr 160
 QY 499 CCCAGGCCGAACTGTCAAGAAGAAGTGAAGCCCAACAGAGCGCCACTTTTCCCTG 558
 Db 161 ProArgProAsnLeuSerLysLysLeuGluAlaProThrGluArgHisPheSerLeu 180
 QY 559 CGAATGAAGACGACCTCACAGCAGAGGCGCGCAGCTCAACCTCAAGAGCGCCACCTGG 618
 Db 181 ArgMetLysSerThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrp 200
 QY 619 AAGGTGCTGCACACTGCTCAGGACATATAGGGCCCTACAGGCCCTTGACACAGACTTCCCT 678
 Db 201 LysValLeuHisCysSerGlyHisMetArgAlaThrLysProProAlaGlnThrSerPro 220
 QY 679 GCCGGAGCCCTCGCTCCGAGCCCTCCCTGCAATGCTGGTGTATCTGTGAAGCCATC 738
 Db 221 AlaGlySerProArgSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIle 240
 QY 739 CCCAGCTCCCTTCCAGATGTTGCTACTCTGGTCTTCCACAGGAGAACTCCCATC 798
 Db 241 ProGlnLeuProPheHisAspGlyAlaThrLeuGlyLeuProGlnGluLysThrProIle 260
 QY 799 TCTACCTTATTACCCCTCTTTGGAAGGCACTACTTTGTCTGTCAAGAGGTGGCTGTT 858
 Db 261 SerThrLeuPheThrProLeuTrpLysAlaLeuLeuCysLeuValLysArgTrpProVal 280
 QY 859 CAGGTGCTACAGGGGAAAGGACTGAATCTCTCTCCCTCATGGGTGTGTGGGCCCTT 918
 Db 281 GlnValLeuGlnGlyLysGlyThrGluSerSerLeuProSerTrpValLeuTrpAlaLeu 300
 QY 919 AACCGGAAAATGTCTCTGCG 939
 Db 301 AsnArgLysAsnCysProGly 307
 RESULT 2
 AAY06295
 ID AAY06295 standard; Protein; 562 AA.
 XX AC AAY06295;
 XX XX
 DT 23-AUG-1999 (first entry)
 DE Mouse transcription regulator MOP7.
 XX KW MOP7; member of the PAS superfamily; bHLH-PAS; mouse;
 KW transcription regulator; hypoxia inducible factor 3 alpha.
 XX OS Mus musculus.
 XX XX WO928464-A2.
 PN 10-JUN-1999.
 PD 27-NOV-1998; 98WO-US25314.
 PF 28-NOV-1997; 97US-0066863.
 XX PR (WISC) WISCONSIN ALUMNI RES FOUND.
 XX PA Bradford CA, Gu YZ, Hogenesch JB;
 PI WPI; 1999-371120/31.
 XX DR N-PSDB; AAX58986.
 DR Developmental signal transduction associated proteins
 XX PT
 XX PS Claim 6; Page 101; 106pp; English.
 XX CC The present sequence represents mouse MOP7, a novel member of the
 CC PAS superfamily, where PAS stands for PER/ARNT/SIM domains. MOP7

CC cDNA (see AAX58986) was identified in a search of murine ESTs designed
 CC to identify bHLH-PAS proteins, and by RACE amplification of lung
 CC cDNA. MOP7 was characterised as hypoxia inducible factor 3 alpha
 CC (HIF 3 alpha). Its expression profile is distinct from that of
 CC HIF 1 alpha (see AAY06289), HIF 2 alpha (see AAY06290), MOP3 (see
 CC AAY06291), Ah receptor and Ah receptor nuclear translocator (ARNT),
 CC suggesting a different functional role. MOP7 probably regulates
 CC the same genes as HIF 1 alpha and 2 alpha, as evidenced by its
 CC dimerisation with the same partners (ARNT, MOP3) and recognition
 CC of the same core response element. MOP7 may have a functional
 CC role associated with response to low oxygen in the tissues in
 CC which it is expressed. The invention provides novel MOPs 2-9
 CC nucleic acids (see AAX58981-88) and proteins (see AAY06289-97).
 CC These are useful in a variety of research, diagnostic and
 CC therapeutic applications. Several of the MOPs are alpha-class
 CC hypoxia-inducible factors. Others are involved in circadian signal
 CC transduction.

XX Sequence 662 AA;

Alignment Scores:
 Pred. No.: 2,03e-85 Length: 662
 Score: 1045.00 Matches: 218
 Percent Similarity: 85.88% Conservative: 1
 Best Local Similarity: 85.49% Mismatches: 5
 Query Match: 51.99% Indels: 31
 DB: 20 Gaps: 3

US-09-896-791B-2 (1-1100) x AAY06295 (1-662)

QY 43 AGGTGCAACCGAGTGGCGGAGGAGAGTCCGGGACGCGCGCGCGCGCGCAGC 102
 Db 7 ArgSerAsnThrGluLeuArgLysGluLysSerArgAspAlaAlaArgSerArgSer 26
 QY 103 CAGGACGAGGAGTGTGTACCGAGCTGGCGGCACACTCTGCCCTTTCGGCGCGCGTACG 162
 Db 27 GlnGluThrGluValLeuTyrGlnLeuAlaHisThrLeuProPheAlaArgGlyValSer 46
 QY 163 GCGCACTGGACAGCGCTCATCATCGCGCTCACAACTACGCTACCTGGCGCATGCAACGC 222
 Db 47 AlaHisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrLeuArgMetHisArg 66
 QY 223 CTCCTGCGCAGCAGTGGAA-----AAAGGGGAGAGCCACTGGACGCGCTG 267
 Db 67 LeuCysAlaAlaGly-GluTrpAsnGlnValGluLysGlyGlyGluProLeuAspAlaC 86
 QY 268 CTACCTGAGGCGCTGGAGGTTTCGTCTGCTACTACCGCGCGAGGAGACATGGCTTA 327
 Db 86 sTyrLeuLysAlaLeuGluGlyPheValMetValLeuThrAlaGluGlyAspMetAla 106
 QY 328 CCTGTCGGAAGTGTGAGCAAGCACCTGGCGCTCAGTCACTGAGTGGACCTCTGTTCTCTCC 387
 Db 106 rLeuSerGluAsnValSerLysHisLeuGlyLeuSerGln----- 119
 QY 388 CTGATACATAACCCCACTCTCTGTGTACCAATTTCTCTCTGAGCTCATTTGGACACAGTATC 447
 Db 120 -----Leu--GluLeuIleGlyHisSerIle 127
 QY 448 TTTGATTTATTCATCCCTGTGTACCAAGAGGAACTTCAGACGCCCTGACCCCGAGCGCG 507
 Db 128 PheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThrProArgPro 147
 QY 508 AACCTGTCAAAGAAAGCTGGAAGCCCAACAGAGCGGCACATTTTCCTCCGCAAGTGAAG 567
 Db 148 AsnLeuSerLysLysLeuGluAlaProThrGluArgHisPheSerLeuArgMetLys 167
 QY 568 AGCAGCTCACAGCAGAGGCGCGCAGCTCAACCTCAAGCGCGCCACCTGGAAAGTGGCTG 627
 Db 168 SerThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrpLysValLeu 187
 QY 628 CACTGCTCAGGACATATGAGGGGCTTCAAGCCCGCTGCACAGACTTCCCTCCCGGAGC 687
 Db 188 HisCysSerGlyHisMetArgAlaTyrLysProAlaGlnThrSerProAlaGlySer 207

QY 688 CCTGCTCGAGCGCTCCCTGCAATGCGTGTCTTATCTGTGAAGCCATCCCC----- 741
 Db 208 ProArgSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIleProHisPro 227
 QY 742 -----CAGTCTCCCTTCCACGATGGTGTACTCTG 771
 Db 228 AlaSerLeuGluProProLeuGlyArgGlyAlaPheLeu 240

RESULT 3

AAB93326

ID AAB93326 standard; Protein; 632 AA.

AC AAB93326;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:12422.

Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

XX EP1074617-A2.

PN 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

DR

XX

PT

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 632 AA;

Alignment Scores:

Pred. No.: 2,08e-82 Length: 632
 Score: 1011.50 Matches: 211
 Percent Similarity: 84.77% Conservatives: 6
 Best Local Similarity: 82.42% Mismatches: 9
 Query Match: 50.32% Indels: 30
 DB: 22 Gaps: 3

US-09-896-791B-2 (1-1100) x AAB93326 (1-632)

Qy 19 ATGGCTGGGGTGCAGCGCGTGGAGTCGAACCCGAGCTGCGGAAGGAGAGTCGCGG 78
 Db 1 MetAlaLeuGlyLeuGlnArgAlaArgSerThrThrGluLeuArgGlyLysSerArg 20
 Qy 79 GACGGCGCCGCGCGCGCGAGCCAGGAGCGAGCGAGTGTGTACCGCTGGCGGCACACT 138
 Db 21 AspAlaAlaArgSerArgSerGlnGlnThrGluValLeuTyrGlnLeuAlaHisThr 40
 Qy 139 CTGCCCTTTGCGCGCGCGTCCAGCGCGCACCTGGACAGGCTCCATCATCGCTCACA 198
 Db 41 LeuProPheAlaArgGlyValSerAlaHisLeuAspLysAlaSerIleMetArgLeuThr 60
 Qy 199 ATCAGCTACCTGGCGATGCACCGCTCTGCGCAGCAGT-----GGAAAA 243
 Db 61 IleSerTyrLeuArgMetHisArgLeuCysAlaAlaGlyGluTrpAsnGlnValGly-Al 80
 Qy 244 AGGGGAGAGCACTGGACGCGCTGCTACCTGAAGCGCTGGAGGTTTCGTATGCTACT 303
 Db 80 aGlyGlyGluProLeuAspAlaCysTyrLeuLysAlaLeuGluGlyPheValMetValLe 100
 Qy 304 CACCGCGGAGGAGACATGCTTACCTGTGCGAAATCTCAGCAAGCACCTGGCGCTCAG 363
 Db 100 uThralaGluGlyAspMetAlaTyrLeuSerGluAsnValSerLysHisLeuGlyLeuSe 120
 Qy 364 TCAGTGGACCTCTGTTCTCTCTCTCTGATACATAACCCCACTCTGTGTACCAATTCTCT 423
 Db 120 rGln-----Le 122
 Qy 424 CTGGAGCTCATTTGACACAGTATCTTTGATTTATCCATCCCTGTGACCAAGAGAACTT 483
 Db 122 u-GluLeuIleGlyHisSerIlePheAspPheIleHisProCysAspGlnGluLeu 141
 Qy 484 CAAGACGCTTACCCCGCGCGCGAACTGTCAAAGAAAGAGCTGGGAAGCCCAACACAG 543
 Db 142 GlnAspAlaLeuThrProGlnGlnThrLeuSerArgArgLysValGluAlaProThrGlu 161
 Qy 544 CGGCACCTTTCCCTGGGAATGAAGACAGCAGCTCACACAGAGGCGCAGCTCAACCTC 603
 Db 162 ArgCysPheSerLeuArgMetLysSerThrLeuThrSerArgGlyArgThrLeuAsnLeu 181
 Qy 604 AAAGCGCCACCTGGAAGTGTCTCACTGCTCAGGACATATGAGGGCTTACAAAGCCCTT 663
 Db 182 LysAlaAlaThrTrpLysValLeuAsnCysSerGlyHisMetArgAlaTyrLysProPro 201
 Qy 664 GCACAGACTTCCCTGCGCGGAGCCCTCGTCCGAGCTCCCTGCAATGCTGGTGTCT 723
 Db 202 ValGlnThrSerProAlaGlySerProAspSerGluProProLeuGlnCysLeuValLeu 221
 Qy 724 ATCTGTGAAGCCATCCCGACGCTCCCTTCCACGATGTGCT 765
 Db 222 IleCysGluAlaIlePro-----HisProGlySer 231

RESULT 4

AAE24222
 ID AAE24222 standard; Protein: 667 AA.

XX AAE24222;

XX AC

XX XX

DT 23-SEP-2002 (first entry)

XX XX

DE Human HIF-3 protein.

XX XX

KW Human; haematologic malignancy; multidrug resistance; MDR; SUMO-1;
 KW hypoxia inducible factor-1; small ubiquitin-like-modifier; HIF-3;
 KW myeloid disorder; chronic lymphoproliferative disorder; lymphoma;
 KW myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;
 KW angiogenic myeloid metaplasia; myeloid leukaemia; gene therapy;
 KW polycythaemia vera; hypoxia responsive element; HRE.

OS Homo sapiens.

XX WO200234291-A2.

XX 02-MAY-2002.

XX 25-OCT-2001; 2001WO-US49856.

XX 26-OCT-2000; 2000US-243542P.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Colgan SP;

XX WPI; 2002-471427/50.

DR N-PSDB; AAD39042.

XX Treating a subject (at risk of) having a hematologic malignancy or
 PT multidrug resistance, e.g. lymphoma or myeloma, by administering
 PT hypoxia inducible factor-1 binding molecules or small
 PT ubiquitin-like-modifier-1 binding molecules -

XX Disclosure; Page 88-91; 92pp; English.

XX The invention relates to a method of treating a subject having or at
 CC risk of developing a haematologic malignancy or multidrug resistance
 CC (MDR). The method involves administering hypoxia inducible factor-1
 CC (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1
 CC binding molecules or HIF-1-SUMO-1 complex modulators. mdrl-hypoxia
 CC responsive element (HRE) binding molecules or antisense nucleic
 CC acid molecules and SUMO-1 binding molecules or antisense molecules
 CC are useful for treating a subject having or at risk of developing
 CC haematologic malignancy or MDR (e.g. a lymphoid disorder or a myeloid
 CC disorder). The lymphoid disorders include lymphocytic leukaemia or
 CC chronic lymphoproliferative disorders e.g. lymphoma, myeloma or chronic
 CC lymphoid leukaemia. The myeloid disorders include chronic or acute
 CC myeloid leukaemia, e.g. angiogenic myeloid metaplasia, essential
 CC thrombocythaemia or polycythaemia vera. The invention is used in gene
 CC therapy. The present sequence is human HIF-3 protein.

XX Sequence 667 AA;

Alignment Scores:

Pred. No.: 8.62e-80 Length: 667
 Score: 982.50 Matches: 205
 Percent Similarity: 85.08% Conservatives: 6
 Best Local Similarity: 82.66% Mismatches: 7
 Query Match: 48.88% Indels: 30
 DB: 23 Gaps: 3

US-09-896-791B-2 (1-1100) x AAE24222 (1-667)

Qy 43 AGGTGCAACCCGAGCTGCGGAAGGAGAGTCCGGGACCGCGCCGCGCGCGCAGC 102
 Db 7 ArgSerThrThrGluLeuArgLysGluLysSerArgAspAlaAlaArgSerArgSer 26
 Qy 103 CAGGAGACGAGGTGTGTACCGAGCTGGCGCACACTCTGCCCTTGGCGGGCGGTGAGC 162
 Db 27 GlnGluThrGluValLeuTyrGlnLeuAlaHisThrLeuProPheAlaArgGlyValSer 46
 Qy 163 GCGCACCTGACAGGCTCCATCATCGCTCAATCAATCAGTACCTGCGCATCGACCGC 222
 Db 47 AlaHisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrLeuArgMetHisArg 66
 Qy 223 CTCTGCGCAGCAGGT-----GGAAAAAGGGGAGAGCCACTGGACGCTG 267
 Db 223 CTCTGCGCAGCAGGT-----GGAAAAAGGGGAGAGCCACTGGACGCTG 267

Db 67 LeuCysAlaAlaGlyGluTrpAsnGlnValGly-AlaGlyGlyGluProLeuAspAlaCys 86
QY 268 CTACCTGAAGCCCTGAGGGTTTCGTGTCATGCTACTACCGCGGAGGAGACATGCTTA 327
Db 86 sTyLeuLysAlaLeuGluGlyPheValMetValLeuThrAlaGluGlyAspMetAlaTy 106
QY 328 CCTGTGGGAAATGTGAGAACACCTGGGCGCTCAGTCAGTCAGGACCTCTGTCTCCTCCTCC 387
Db 106 rLeuSerGluAsnValSerLysHisLeuGlyLeuSerGln----- 119
QY 388 CTGATACATAACCCCACTCCTGGTACCAATTTCTCTCTGGAGCTATTGGACACAGTATC 447
Db 120 -----Leu--GluLeuIleGlyHisSerIle 127
QY 448 TTTGATTTTATCATCCTCTGACCAAGAGAACTTCAAGACGCCCTGACCCCGAGCGCG 507
Db 128 PheAspPheIleHisProCysAspGlnGluLeuGlnAspAlaLeuThrProGlnGln 147
QY 508 AACCTGTCAAGAAAGCTGGAAGCCCAAGAGCGCCACTTTTCCCTGCGAATGAAG 567
Db 148 ThrLeuSerArgArgLysValGluAlaProThrGluArgCysPheSerLeuArgMetLys 167
QY 568 AGACGCTCACAGACAGAGCGCGAGCTCAACCTCAAGCGGCCACCTGGAGGTCGTG 627
Db 168 SerThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrpLysValLeu 187
QY 628 CACTGCTCAGGACATATGAGGCGCTACAGCCCTGACAGACTTCCTCCGCGGAGC 687
Db 188 AsnCysSerGlyHisMetArgAlaTyLysProAlaGlnThrSerProAlaGlySer 207
QY 688 CTTGCTCGGAGCCTCCCTGCAATGCTGCTGCTTATCTGTGAAGCCCATCCCGAGCTC 747
Db 208 ProAspSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIlePro----- 225
QY 748 CCCTTCCAGATGGTGCT 765
Db 226 -----HisProGlySer 229
RESULT 5
ABG66737
ID ABG66737 standard; Protein; 407 AA.
XX AC ABG66737;
XX DT 30-AUG-2002 (first entry)
XX DE Human novel polypeptide #72.
XX KW Human; inflammatory condition; shock; sepsis; immune response;
KW cancer; wound healing; central nervous system disease; haematopoiesis;
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
KW bone degenerative disorder; periodontal disease; reperfusion injury;
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
KW allergic condition; thrombolysis; thrombosis; coagulation disorder;
KW fungal infection.
XX OS Homo sapiens.
XX PN WO200244340-A2.
XX PD 06-JUN-2002.
XX PF 30-NOV-2001; 2001WO-US47004.
XX PR 30-NOV-2000; 2000US-0028952.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
PI Yamazaki V, Ujwal ML, Drmanac RT;
XX

DR WPI; 2002-508509/54.
XX N-PSDB; ABK94961.
PT Novel nucleic acids and polypeptides for diagnosis, treatment of
PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
PT disorders, cancer and promoting wound healing.
XX Claim 10; Page 649; 672pp; English.
XX The invention relates to human novel polynucleotides and associated
CC polypeptides. The polynucleotides and polypeptides are useful for
CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
CC and cancer and for promoting wound healing. The sequences are used to
CC induce the proliferation of neural cells and regeneration of nerve and
CC brain tissue, and are useful for the treatment of central and peripheral
CC Parkinson's disease, Huntington's disease and amyotrophic lateral
CC sclerosis. The sequences are involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid
CC cell disorders and platelet disorders such as thrombocytopenia, tissue
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis
CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human
CC novel polypeptides of the invention.
XX SQ Sequence 407 AA:

Alignment Scores:
Pred. No.: 1.89e-79 Length: 407
Score: 978.00 Matches: 229
Percent Similarity: 67.23% Conservative: 11
Best Local Similarity: 64.15% Mismatches: 37
Query Match: 48.66% Indels: 80
DB: 23 Gaps: 10

US-09-896-791B-2 (1-1100) x ABG66737 (1-407)

QY 46 TCGAACACGAGCTGCGGAGGAGAGACTCGCGGGAGCGCGCGGAGCGCGGAGCCAG 105
Db 54 SerThrThrGluLeuArgLysGluLysSerArgAspAlaAlaArgSerArgSerGln 73
QY 106 GAGACGGAGGTGCTGTACCAGCTGGCGGCACACTCTCCCTTTTGGCGCGCGGTCAGCGC 165
Db 74 GluThrGluValLeuTyrglnLeuAlaHisThrLeuProPheAlaargGlyValSerAla 93
QY 166 CACCTGGACAAGCGCTCCCATCATGCGCCCTCACAAATCAGTACCTGCGCATGCCGCCCTC 225
Db 94 HisLeuAspLysAlaSerIleMetArgLeuThrIleSerTyrrLeuArgMetHisArgLeu 113
QY 226 TCGCGAGCAGGT-----GGAAAAGGGGAGAGCCCACTGGACGCCCTGCTA 270
Db 114 CysAlaAlaGlyGluTrpAsnGlnValGly-AlaGlyGlyGluProLeuAspAlaCyst 133
QY 271 CCTGAGGCGCCTGGAGGGTTTCGTGTCATGCTACTCACCGCGGAGGAGACATGCTTACCT 330
Db 133 rLeuLysAlaLeuGluGlyPheValMetValLeuThrAlaGluGlyAspMetAlaTyrl 153
QY 331 GTCGAAAAATGTCAGCAAGCACCTGGGCGCTCAGTCAGTGGAGCCTCTGTCTCTCCTCCCTG 390
Db 153 uSerGluAsnValSerLysHisLeuGlyLeuSerGln----- 165
QY 391 ATACATAACCCCACTCCTGGTACCAATTTCTCTCTGAGCTCATTTGGACACAGTATCTTT 450
Db 166 -----Leu--GluLeuIleGlyHisSerIlePhe 174

```
QY 451 GATTTTATCCATCCTGTGACCAAGAGAACTTCAAGACGCCCTGACCCCGAGCCGCAAC 510
Db 175 AspheIleHisProCysAspGlnGluGluLeuGlnAspAlaLeuThrProGlnGlnThr 194
QY 511 CTGTCAAGAGAGAGCTGGAGCCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGAGC 570
Db 195 LeuSerArgArgLysValGluAlaProThrGluArgCysPheSerLeuArgMetLysSer 214
QY 571 ACGCTCACAGCAGAGGCGGCACCTCAACCTCAAGCGGCCACTGGAAGTGCTGCAC 630
Db 215 ThrLeuThrSerArgGlyArgThrLeuAsnLeuLysAlaAlaThrTrpLysValLeuAsn 234
QY 631 TGCTCAGACATATGAGGGCTACAGCCGCCCTGCACAGACTTCCCTGCGGGAGCCCT 690
Db 235 CysSerGlyHisMetArgAlaTyrLysProProAlaGlnThrSerProAlaGlySerPro 254
QY 691 CGCTCCGAGCTCCCTGCAATGCTGGTCTTATCTGTGAGCCATCCCCAGCTCCCC 750
Db 255 AspSerGluProProLeuGlnCysLeuValLeuIleCysGluAlaIlePro 271
QY 751 TTCACGATGGTGTACTCTGGGTCTTCCACAGGAGAAGACTCCCATCTCTACCTTATTC 810
Db 272 ---HisProGly-----SerLeuGlu 277
QY 811 ACCCTCTTTGGAAGCAGCTACTTTGTCTGTCAAGAGGTGGCTG----- 856
Db 278 ProProLeu-GlyArgGlyAlaPhe---LeuSerArgHisSerLeuAspMetLysPheTh 296
QY 857 -----TTCAGGTGCTACAGGGGAAGGAGTCAATCT----- 889
Db 296 rTyrCysAspAspArgIleAlaGluValAlaGlyTyrSerProAspLeuIleGlyCy 316
QY 890 -----CTCTCCCTCATGG 903
Db 316 sSerAlaTyrGluTyrIleHisAlaLeuAspSerAspAlaValSerLysSerIleHisTh 336
QY 904 GTGTGTGGGCCCTTAACCGGAAAAATGTCTGGCTAGGAGGAGTGCAAGACATGCC 963
Db 336 rCysMetTyrProIleSerPro-----GlyAlaLysProAlaAlaThrTrpPr 352
QY 964 CAGCTATCTTAGCCCAACACCCACAATGTCTCGAAACCA 1006
Db 352 oProAlaAsp---ThrArgThrProGlnLeu---ProIlePro 364

RESULT 6
AAB94934
ID AAB94934 standard; Protein: 648 AA.
XX AC
XX DT
XX TT
XX DE
XX DE
XX KW
XX OS
XX XX
XX PN
XX XX
XX PD
XX PF
XX PR
XX PR
XX PR
XX PR
XX PR
XX PA
XX XX
XX PI
```

```
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX PS Claim 8; SEQ ID 16415; 2537pp + CD ROM; English.
XX
XX CC The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX SQ Sequence 648 AA;

Alignment Scores:
Pred. No.: 3,56e-59 Length: 648
Score: 753.50 Matches: 158
Percent Similarity: 82.00% Conservative: 6
Best Local Similarity: 79.00% Mismatches: 6
Query Match: 37.49% Indels: 30
DB: 22 Gaps: 3

US-09-896-791B-2 (1-1100) x AAB94934 (1-648)
QY 187 ATGGCGCTCAATCAGCTACCTACCTGCGCATGCACGCCCTCTGCGCAGCAGGT----- 237
Db 1 MetArgLeuThrIleSerTyrLeuArgMetHisArgLeuCysAlaAlaGlyGluTrpAsn 20
QY 238 -----GGAAAAAGGGGAGAGCCACTGGAGCGCTGCTACCTGAGGCCCTGGAGGGTTT 291
Db 21 GlnValGly-AlaGlyGlyGluProLeuAspAlaCysTyrLeuLysAlaLeuGluGlyPh 40
QY 292 CGTCATGGTACTCACCGCCGAGGAGACATGGCTTACCTGCGAAAATGTCAGCAAGCA 351
Db 40 eValMetValLeuThrAlaGluGlyAspMetAlaTyrLeuSerGluAsnValSerLysHi 60
QY 352 CTGTGGGCGCTCAGTCAGTGGACCTCTGTTCCTCCTCCTGATACATACCCACCTCCTGGT 411
Db 60 sLeuGlyLeuSerGln----- 65
QY 412 ACCAATTTCTCTGAGAGCTCATTGGACAGCATATCTTTGATTATTCATCCCTGTGAC 471
Db 66 -----Leu--GluLeuIleGlyHisSerIlePheAspPheIleHisProCysAsp 81
QY 472 CAAGAGAGAACTTCAAGACGCCCTGACCCCGAGCGCAACCTGTCAAAGAGAAAGCTGGA 531
Db 82 GlnGluGluLeuGlnAspAlaLeuThrProGlnGlnThrLeuSerArgArgLysValGlu 101
QY 532 GCCCAACAGAGCGCCACTTTTCCCTGCGAATGAAGAGACGCTCACCACAGAGCGCC 591
Db 102 AlaProThrGluArgCysPheSerLeuArgMetLysSerThrLeuThrSerArgGlyArg 121
```


CC The invention relates to isolated variants of HIF-1alpha, such as that
CC represented by the present sequence. The variants are useful for
CC identifying compounds capable of modulating the function of a functional
CC domain of human HIF-1alpha. The method comprises contacting a candidate
CC compound with a cell expressing a HIF-1alpha variant conjugated to a
CC molecular probe. The localization of the probe can be detected in the
CC cell. The Aequorea victoria green fluorescent protein can be used as the
CC molecular probe. The compounds are useful for the regulation of
CC HIF-1alpha target genes, such as those involved in the regulation of
CC angiogenesis, erythropoiesis and glycolysis.

XX Sequence 245 AA;

Alignment Scores:
Pred. No.: 4-57e-46 Length: 245
Score: 606.50 Matches: 129
Percent Similarity: 58.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 21 Gaps: 3

US-09-896-791B-2 (1-1100) x AAY94627 (1-245)

QY 49 AACACCGAGCTCGGAAGAGAGTTCGGCGGAGCGCGCGCGCGCGCGCGCGAGG 108
Db 14 SerSerGluArgArgLysGluSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGGAGGTCTTACCAGCTGGCGCACACTCTGCCCTTTGGCGGCGCTCAGCGGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerHis 53
QY 169 CTGGACAGGCTCCATCATCTGCGCTCACAATCAGTCTGCGTACGCGACCGCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeu 73
QY 229 GCAGCAGGTGGAAAGGGGGAGAG-----CCACTGGAGCGCTGTACTAC 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGCGCTGGAGGTTCTGCTCATCTGCTACTCACCCTGCGGAGGAGACATGCTTCTGTC 333
Db 93 uIyAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTCAGACAGCACTGGCGCTCAGTCTGAGTGGAGCTCTGTCTCTCTCCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATACCCCACTCTCTGGTACCAATTTCTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCATCTCTGTGACCAAGAGNACTTCAAGACGCGCTGACCCCGAGGCGGAACTTG 513
Db 135 PheThrHisProCysAspHisGluGluMetArgGlnMetLeuThrHisArgAsnGlyLeu 154
514 TCAAGAGAAGCTGAAGCCCAACAGAGCGCGCACTTTTCTCGCAATGAAGACAGCAG 573
Db 155 ValLysLysGlyGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACAGCAGAGGCGCGCGCTCAACCTCAAGCGCGCCACCTGGAAGGTGTGCTGCACTGC 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATAGGCGCTTACAGCCCTCGCACAGACTTCCCTGCGCGGAGCCCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCTCTCCCTGCAATGCTGTGCTTATCTGTGAAGCCATCCCC 741
Db 213 LysLysProProMetThrCysLeuValLeuIleCysGluProIlePro 228
RESULT 10
AAY94628

ID AAY94628 standard; protein; 330 AA.
XX
AC AAY94628;
XX
DT 15-AUG-2000 (first entry)
XX
-DE HIF-1alpha variant protein sequence HIF-1alpha/1-330.
KW Hypoxia-inducible factor 1alpha; HIF-1alpha; PAS-B; N-TAD; C-TAD;
KW regulation; angiogenesis; erythropoiesis; glycolysis; human.
XX
OS Homo sapiens.
XX
PN WO200029437-A1.
XX
PD 25-MAY-2000.
XX
PF 11-NOV-1999; 99WO-SE02053.
XX
PR 13-NOV-1998; 98SE-0003891.
XX
PA (PHAA) PHARMACIA & UPJOHN AB.
XX
PI Berkenstam A, Poellinger L;
XX
DR WPI; 2000-399715/34.
XX
PT Human hypoxia-inducible factor alpha variants for identifying compounds
PT that modulate its functional domain and regulate genes involved in
PT angiogenesis, erythropoiesis -
XX
PS Claim 15; Page 69-70; 87pp; English.

CC This sequence represents a fragment of the hypoxia-inducible factor
CC (HIF)-1alpha amino acid sequence. The mechanism of action of HIF-1alpha
CC is a multi-step process which includes hypoxia-dependent nuclear import
CC and activation of the transactivation domain. The HIF-1alpha consists of
CC a number of functional domains including a PAS-B (Per, Arnt, Sim) domain
CC located in human HIF-1alpha between amino acids 173 and 390, a C-terminal
CC nuclear localization sequence located at amino acids 718-584, a
CC transactivator domain (N-TAD) located between amino acids 531 and 584,
CC and a second transactivator domain (C-TAD) located between 813 and 826.
CC The invention relates to isolated variants of HIF-1alpha, such as that
CC represented by the present sequence. The variants are useful for
CC identifying compounds capable of modulating the function of a functional
CC domain of human HIF-1alpha. The method comprises contacting a candidate
CC compound with a cell expressing a HIF-1alpha variant conjugated to a
CC molecular probe. The localization of the probe can be detected in the
CC cell. The Aequorea victoria green fluorescent protein can be used as the
CC molecular probe. The compounds are useful for the regulation of
CC HIF-1alpha target genes, such as those involved in the regulation of
CC angiogenesis, erythropoiesis and glycolysis.

XX SQ Sequence 330 AA;

Alignment Scores:
Pred. No.: 5e-46 Length: 330
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 21 Gaps: 3

US-09-896-791B-2 (1-1100) x AAY94628 (1-330)

QY 49 AACACCGAGCTCGGAAGAGAGTTCGGCGGAGCGCGCGCGCGCGCGAGG 108
Db 14 SerSerGluArgArgLysGluSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGGAGGTCTGTACCAGCTGGCGCACACTCTGCCCTTTGGCGGCGCTCAGCGGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerHis 53

PF 25-AUG-1999; 99WO-US194116.
XX
PR
XX 25-AUG-1998; 98US-0148547.
XX
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
XX Semenza GL;
XX
DR WPI; 2000-246493/21.
XX
XX Variant forms of hypoxia-inducible factor (HIF)-1 alpha, useful for
PT treating hypoxia or ischemia-related tissue damage -
XX
PS
XX Claim 1; Page -; 96pp; English.
-XX
CC The present sequence represents a variant of hypoxia-inducible factor
CC (HIF)-1 alpha, comprising amino acids 1-391 and 549-826 of the wild
CC type protein (see AAY69407). The HIF-1alpha variants are stable under
CC hypoxic and non-hypoxic conditions. The variants comprises amino acid
CC residues 1-391 and 521-826, 549-826, 576-826, 429-826, 469-826, 494-826,
CC 508-826, 512-826 or 517-826 of the wild type human HIF-1alpha
CC polypeptide, in which residues 551 and 552 are not serine and threonine,
CC respectively. The HIF-1alpha variant polynucleotide sequences are useful
CC for increasing expression of a hypoxia inducible gene in a cell. They
CC is also useful for providing constitutive expression of a hypoxia
CC inducible factor in a cell, and for reducing or preventing hypoxia or
CC ischemia related damage. The variant HIF-1alpha polypeptides are useful
CC for providing prophylactic therapy for inducing the level of
CC angiogenesis in tissues of patients at risk of coronary artery disease
CC or ischemic tissue damage.
CC note: this sequence does not appear in the specification; it was created
CC using information provided.
XX
SQ Sequence 669 AA;

Alignment Scores:
Pred. No.: 6,2e-46 Length: 669
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservative: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 21 Gaps: 3

US-09-896-791B-2 (1-1100) x AAY84167 (1-669)

QY 49 AACACCGAGCTGCGAGGAGGAAGTTCGGGACGCGCGCGCGCGCGCGAGGAG 108
DB 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGAGGTGTGTACCGAGTGGCGCACACTCTGCGCTTTGCGCGCGCGCGCGCGCAC 168
DB 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerSerHis 53
QY 169 CTGACAAAGGCTCCATCAGTCCGCTCAGATCAGTACCTGCGCATGCGACCGCTCTGC 228
DB 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeu 73
QY 229 GCACGAGTGGAAAAGGGGAGAG-----CCACTGGACGCTGCTACCT 273
DB 74 AspAlaGlyAspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCCCTGGAGGGTTTCGTATGGTACTACCCGCGGAGGAGACATGGCTTACCTGTC 333
DB 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTCAGCAAGCACCTGGCGCTCAGTCAGTGGACCTCTGCTCCTCCTCCTGATA 393
DB 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAACCCCACTCCTGGTACCATTCTCTCTGAGCTCATTCGACACACTACTTTGAT 453
DB 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134

QY 454 TTTATCATCCCTGTGACCAAGAGGAAGTTCACAGACGCCCTGACCCCGAGCCGGAACCTG 513
DB 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAAGCTGGGAAGCCCAACAGAGCGGCACCTTTCCCTGCGCAATGAAGAGCAG 573
DB 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPheLeuArgMetLysCysThr 174
QY 574 CTCACGACAGAGCGCGCACGCTCAACCTCAAGCGCGCACCTGGAAGGTGCTGCTGCTG 633
DB 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrIleLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGGCTACAGCCCGCTGCACAGACTTCCCTCGCGGAGCCCTCGC 693
DB 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCTCCCTGCGCAATGCTGCTGCTTATCTGTGCAAGCATCCCT 741
DB 213 LysLysProMetThrCysLeuValLeuLysGluProIlePro 228
RESULT 14
AAY84166
ID AAY84166 standard; Protein; 697 AA.
XX
AC AAY84166;
XX
DT 03-JUL-2000 (first entry)
XX
DE A variant of human hypoxia inducible factor-1 alpha protein.
XX
KW Human; hypoxia-inducible factor 1 alpha; HIF-1alpha; variant;
KW hypoxia inducible gene; hypoxia inducible factor; hypoxia;
KW ischemia related damage; angiogenesis; coronary artery disease;
KW ischemic tissue damage.
XX
OS Synthetic.
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Misc-difference 422
FT /note= "this residue is optionally not Ser, and is
FT Misc-difference 423 preferably Gly"
FT /note= "this residue is optionally not Thr, and is
FT preferably Ala"
XX
PN WO200010578-A1.
XX
PD 02-MAR-2000.
XX
XX 25-AUG-1999; 99WO-US194116.
XX
XX 25-AUG-1998; 98US-0148547.
XX
XX (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
XX Semenza GL;
XX
XX WPI; 2000-246493/21.
XX
XX Variant forms of hypoxia-inducible factor (HIF)-1 alpha, useful for
PT treating hypoxia or ischemia-related tissue damage -
XX
XX Claim 1; Page -; 96pp; English.
XX
CC The present sequence represents a variant of hypoxia-inducible factor
CC (HIF)-1 alpha, comprising amino acids 1-391 and 521-826 of the wild
CC type protein (see AAY69407). The HIF-1alpha variants are stable under
CC hypoxic and non-hypoxic conditions. The variants comprises amino acid
CC residues 1-391 and 521-826, 549-826, 576-826, 429-826, 469-826, 494-826,
CC 508-826, 512-826 or 517-826 of the wild type human HIF-1alpha
CC polypeptide, in which residues 551 and 552 are not serine and threonine,
CC respectively. The HIF-1alpha variant polynucleotide sequences are useful
CC

CC for increasing expression of a hypoxia inducible gene in a cell. They
CC is also useful for providing constitutive expression of a hypoxia
CC inducible factor in a cell, and for reducing or preventing hypoxia or
CC ischemia related damage. The variant HIF-1alpha polypeptides are useful
CC for providing prophylactic therapy for inducing the level of
CC angiogenesis in tissues of patients at risk of coronary artery disease
CC or ischemic tissue damage.
CC note: this sequence does not appear in the specification; it was created
CC using information provided.
XX Sequence 697 AA;

Alignment Scores:
Pred. No.: 6,28e-46 Length: 697
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservatives: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 21 Gaps: 3

US-09-896-791B-2 (1-1100) x AAY84166 (1-697)

QY 49 AACACCGAGTCGGGAAGAGAGAGTCCGGCGACGCGCGCGCGCGCGCGCGCGCGAGGAG 108
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGAGGTGCTGCTACAGCTGGCGGCACACTCTGCCCTTTTCGCGCGCGCGCGCGCGAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProHisAsnValSerSerHis 53
QY 169 CTGACAAGGCTCCATCATGCGCCCTCACAACTACGTCGTCGTCGTCGTCGTCGTCGTCG 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
QY 229 GCAGAGGTGGAAAAGGGGGAGAG-----CCACTGGAGCGCTGCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCCCTGGAGGTTTCTGCTACTCTACTCACCAGCGGGGAGACATGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GAAAAATGTCAAGAACGCTGGCGCTCAGTCAGTCAGTCAGTCGTCCTCTCTCTCTCTCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAACCCCACTCCTGGTACCAATTTCTCTCTGAGCTCATGTCGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCATCCCTGTGACCAAGAGAACTTCAAGAGCCGCTGACCCCGCGCGCGCGCGCGCG 513
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAAGCTGGAAGCCCAACAGAGCCGCTCTTTCCTGCGAATGAAGAGCACG 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACAGCAGAGGGCGCGCTGCTCAACCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGCGCTTACAGCGCCCTGCGACACACTTCCCTGCGCGCGCGCGCGCG 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCCTCCCTGCAATGCTGGTGTATCTGTAAGCCATCCCG 741
Db 213 LysLysProPromethrCysLeuValLeuIleCysGluProIlePro 228

RESULT 15
AAY84173
ID AAY84173 standard; Protein; 701 AA.

XX AAY84173;
XX 03-JUL-2000 (first entry)
XX A variant of human hypoxia inducible factor-1 alpha protein.
XX Human; hypoxia-inducible factor 1 alpha; HIF-1alpha; variant;
KW hypoxia inducible gene; hypoxia inducible factor; hypoxia;
KW ischemia related damage; angiogenesis; coronary artery disease;
KW ischemic tissue damage.
XX Synthetic.
OS Homo sapiens.
XX Key Location/Qualifiers
FT Misc-difference 426 /note= "this residue is not Ser, and is
FT preferably Gly"
FT Misc-difference 427 /note= "this residue is not Thr, and is
FT preferably Ala"
XX WO200010578-A1.
XX 02-MAR-2000.
XX 25-AUG-1999; 99WO-US19416.
XX 25-AUG-1998; 98US-0148547.
XX (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX Semenza GL;
XX WPI; 2000-246493/21.
XX Variant forms of hypoxia-inducible factor (HIF)-1 alpha, useful for
XX treating hypoxia or ischemia-related tissue damage -
XX Claim 1; Page -; 96pp; English.

XX The present sequence represents a variant of hypoxia-inducible factor
XX (HIF)-1 alpha, comprising amino acids 1-391 and 517-826 of the wild
XX type protein (see AAY69407). The HIF-1alpha variants are stable under
XX hypoxic and non-hypoxic conditions. The variants comprises amino acid
XX residues 1-391 and 521-826, 576-826, 429-826, 469-826, 494-826,
XX 508-826, 512-826 or 517-826 of the wild type human HIF-1alpha
XX polypeptide, in which residues 551 and 552 are not serine and threonine,
XX respectively. The HIF-1alpha variant polynucleotide sequences are useful
XX for increasing expression of a hypoxia inducible gene in a cell. They
XX is also useful for providing constitutive expression of a hypoxia
XX inducible factor in a cell, and for reducing or preventing hypoxia or
XX ischemia related damage. The variant HIF-1alpha polypeptides are useful
XX for providing prophylactic therapy for inducing the level of
XX angiogenesis in tissues of patients at risk of coronary artery disease
XX or ischemic tissue damage.
XX note: this sequence does not appear in the specification; it was created
XX using information provided.

XX Sequence 701 AA;
Alignment Scores:
Pred. No.: 6,29e-46 Length: 701
Score: 606.50 Matches: 129
Percent Similarity: 68.07% Conservatives: 33
Best Local Similarity: 54.20% Mismatches: 48
Query Match: 30.17% Indels: 28
DB: 21 Gaps: 3

US-09-896-791B-2 (1-1100) x AAY84173 (1-701)

QY 49 AACACCGAGTCGGGAAGAGAGTCCGGCGACGCGCGCGCGCGCGCGCGCGAGGAG 108

```
Db 14 SerSerGluArgArgLysGluLysSerArgAspAlaAlaArgSerArgSerLysGlu 33
QY 109 ACGGAGGTGCTGTACACGCGGCACACTGCGCTTGGCGCGCGCTCAGCGCGCAC 168
Db 34 SerGluValPheTyrGluLeuAlaHisGlnLeuProLeuProHisAsnValSerHis 53
QY 169 CTGACAAAGGCTCCATCATCGCGCTCACAAATCAGCTACCTGCGCATGCGACCGCTCTGC 228
Db 54 LeuAspLysAlaSerValMetArgLeuThrIleSerTyrLeuArgValArgLysLeuLeu 73
QY 229 GCACAGGTGGAAAAGGGGAGAG-----CCACTGGACGCGCTGCTACCT 273
Db 74 AspAlaGly-AspLeuAspIleGluAspAspMetLysAlaGlnMetAsnCysPheTyrLe 93
QY 274 GAAGGCGCTGGAGGTTTCGTCTGCTGACTCACCGCGCGAGGAGACATGGCTTACCTGTC 333
Db 93 uLysAlaLeuAspGlyPheValMetValLeuThrAspAspGlyAspMetIleTyrIleSe 113
QY 334 GGAATATGTCAGCAAGCACCTGGGCGCTCAGTCACTGAGTGGACCTCTGCTCCTCCTGATA 393
Db 113 rAspAsnValAsnLysTyrMetGlyLeuThrGln----- 124
QY 394 CATAAACCCCACTCCTGTTACCAATTCTCTCTGGAGCTCATTTGGACACAGTATCTTTGAT 453
Db 125 -----Phe--GluLeuThrGlyHisSerValPheAsp 134
QY 454 TTTATCCATCCCTGTGACCAAGAGAACTTCAAGACGCGCTGACCCCGCGCGAACCTG 513
Db 135 PheThrHisProCysAspHisGluGluMetArgGluMetLeuThrHisArgAsnGlyLeu 154
QY 514 TCAAGAAGAGAGCTGGAAGCCCAACACAGAGCGCACCTTTCCCTGCGAATGAAGAGCACG 573
Db 155 ValLysLysGlyLysGluGlnAsnThrGlnArgSerPhePheLeuArgMetLysCysThr 174
QY 574 CTCACAGCAGAGGGCGCACGCTCAACCTCAAAGCGGCCACCTGGAAGGTGCTGCACTGC 633
Db 175 LeuThrSerArgGlyArgThrMetAsnIleLysSerAlaThrTrpLysValLeuHisCys 194
QY 634 TCAGGACATATGAGGCGCTACAGCCCTGACAGACCTTCCCTGCGGGAGGCCCTCGC 693
Db 195 ThrGlyHisIleHisValTyrAspThrAsnSerAsnGlnProGlnCysGly-----Tyr 212
QY 694 TCCGAGCTCCCTGCAATGCTGGTCTTATCTGTGAGGCCATCCCC 741
Db 213 LysLysProPrometThrCysLeuValLeuIleCysGluProIlePro 228
```

Search completed: December 2, 2002, 20:36:01
Job time : 63.5 secs